目 录

《俄罗斯数学教材选译》序

第 4 版和第 3 版序言

第 2 版序言

第 1 版序言摘录

第一章 一些通用的数学概念与记号 1

§1. 逻辑符号 1

1. 关系与括号 (1) 2. 关于证明的注记 (2) 3. 某些专门记号 (3) 4. 最后的注记 (3) 练习 (3)

§2. 集与集的初等运算 4

1. 集合的概念 (4) 2. 包含关系 (5) 3. 最简单的集合运算 (7) 练习 (9)

§3. 函数 10

1. 函数 (映射) 的概念 (10) 2. 映射的简单分类 (13) 3. 函数的复合与互逆映射 (14) 4. 作为关系的函数. 函数的图像 (16) 练习 (19)

§4. 某些补充 21

1. 集的势 (基数) (21) 2. 公理化集合论 (23) 3. 关于数学命题的逻辑结构及其用集合论语言的写法的注记 (24) 练习 (26)

第二章 实数 29

§1. 实数集的公理系统及它的某些一般性质 29

1. 实数集的定义 (29) 2. 实数的某些一般的代数性质 (32) 3. 完备公理与数集的上 (下) 确界的存在性 (36)

§2. 最重要的实数类及实数计算方面的一些问题 38

1. 自然数与数学归纳原理 (38) 2. 有理数与无理数 (40) 3. 阿基米德原理 (43) 4. 实数集的几何解释与实数计算方面的一些问题 (44) 练习 (54)

§3. 与实数集的完备性有关的基本引理 58

1. 闭区间套引理 (柯西-康托尔原理) (58) 2. 有限覆盖引理 (博雷尔-勒贝格原理) (59) 3. 极限点引理 (波尔察诺-魏尔斯特拉斯原理) (60) 练习 (60)

§4. 可数集与不可数集 61

1. 可数集 (61) 2. 连续统的势 (63) 练习 (63)

第三章 极限 66

§1. 序列的极限 66

1. 定义和例子 (66) 2. 数列极限的性质 (68) 3. 数列极限的存在问题 (72) 4. 级数的初步知识 (81) 练习 (90)

§2. 函数的极限 93

1. 定义和例子 (93) 2. 函数极限的性质 (97) 3. 函数极限的一般定义 (对基的极限) (112) 4. 函数极限的存在问题 (115) 练习 (130)

第四章 连续函数 133

§1. 基本定义和例子 133

1. 函数在一点处的连续性 (133) 2. 间断点 (137)

§2. 连续函数的性质 140

1. 局部性质 (140) 2. 连续函数的整体性质 (141) 练习 (149)

第五章 微分学 154

§1. 可微函数 154

1. 问题和引言 (154) 2. 在一点处可微的函数 (158) 3. 切线; 导数和微分的几何意义 (161) 4. 坐标系的作用 (163) 5. 一些例子 (165) 练习 (170)

§2. 微分的基本法则 171

1. 微分法和算术运算 (171) 2. 复合函数的微分法 (175) 3. 反函数的微分法 (178) 4. 基本初等函数的导数表 (182) 5. 最简单的隐函数的微分法 (183) 6. 高阶导数 (187) 练习 (190)

§3. 微分学的基本定理 191

1. 费马引理和罗尔定理 (191) 2. 拉格朗日和柯西的有限增量定理 (193) 3. 泰勒公式 (196) 练习 (208)

§4. 用微分学的方法研究函数 212

1. 函数单调的条件(参看函数单调性检验法) (212) 2. 函数的内极值点条件 (213) 3. 函数凸的条件 (218) 4. 洛必达法则 (225) 5. 作函数的图像 (226) 练习 (235)

§5. 复数 初等函数彼此间的联系 238

1. 复数 (238) 2. C 中的收敛及复数项级数 (241) 3. 欧拉公式以及初等函数彼此间的联系 (245) 4. 函数的幂级数表示, 解析性 (248) 5. 复数域 C 的代数封闭性 (253) 练习 (259)

§6. 自然科学中应用微分学的一些例子 260

1. 齐奥尔科夫斯基公式 (260) 2. 气压公式 (262) 3. 放射衰变、连锁反应及原子反应堆 (264) 4. 空气中的落体 (266) 5. 再谈数 e 及指数函数 $\exp x$ (267) 6. 振动 (270) 练习 (273)

§7. 原函数 276

1. 原函数和不定积分 (276) 2. 求原函数的基本的一般方法 (278) 3. 有理函数的原函数 (284) 4. $\int R(\cos x, \sin x)dx$ 型的原函数 (287) 5. $\int R(x, y(x))dx$ 型的原函数 (289) 练习 (292)

第六章 积分 298

§1. 积分定义和可积函数集的描述 298

1. 问题和启发性想法 (298) 2. 黎曼积分的定义 (299) 3. 可积函数集 (301) 练习 (312)

§2. 积分的线性性、可加性和单调性 314

1. 作为空间 $\mathcal{R}[a,b]$ 上的线性函数的积分 (314) 2. 作为积分区间的可加函数的积分 (314) 3. 积分的估计, 积分的单调性和中值定理 (316) 练习 (323)

§3. 积分和导数 324

1. 积分和原函数 (324) 2. 牛顿–莱布尼茨公式 (326) 3. 定积分的分部积分法和泰勒公式 (326) 4. 定积分中的变量替换 (328) 5. 一些例子 (330) 练习 (334)

§4. 积分的一些应用 336

1. 定向区间的可加函数和积分 (336) 2. 道路的长度 (338) 3. 曲边梯形的面积 (343) 4. 旋转体的体积 (344) 5. 功与能 (345) 练习 (350)

§5. 反常积分 351

1. 反常积分的定义、例题和基本性质 (351) 2. 反常积分收敛性的研究 (355) 3. 具有几个奇异点的反常积分 (360) 练习 (362)

第七章　多变量函数和它的极限与连续性 365

§1. 空间 $\mathbb{R}^m$ 和它的重要子集类 365

1. 集合 $\mathbb{R}^m$ 和 $\mathbb{R}^m$ 中的距离 (365) 2. $\mathbb{R}^m$ 中的开集与闭集 (367) 3. $\mathbb{R}^m$ 中的紧集 (369) 练习 (371)

§2. 多变量函数的极限与连续性 371

1. 函数的极限 (371) 2. 多变量连续函数及其性质 (376) 练习 (380)

第八章　多变量函数微分学 381

§1. $\mathbb{R}^m$ 中的线性结构 381

1. 作为向量空间的 $\mathbb{R}^m$ (381) 2. 线性映射 (382) 3. $\mathbb{R}^m$ 中的范数. (383) 4. $\mathbb{R}^m$ 的欧几里得结构 (384)

§2. 多变量函数的微分 386

1. 多变量可微函数和函数在一点的微分 (386) 2. 实值函数的偏导数与微分 (387) 3. 映射的微分的坐标表示. 雅可比矩阵 (389) 4. 函数在一点的连续性、偏导数和可微性 (390)

§3. 微分法的基本定律 391

1. 微分法运算的线性性质 (391) 2. 复合映射的微分法 (393) 3. 逆映射的微分法 (397) 练习 (399)

§4. 多变量实值函数微分学的基本事实 403

1. 中值定理 (403) 2. 多变量函数可微性的充分条件 (405) 3. 高阶偏导数 (406) 4. 泰勒公式 (409) 5. 多变量函数的极值 (411) 6. 与多变量函数有关的某些几何形象 (417) 练习 (421)

§5. 隐函数定理 426

1. 问题的提出与启发性想法 (426) 2. 隐函数定理的最简单情形 (428) 3. 过渡到依赖关系 $F(x^1,\cdots,x^m,y)=0$ 的情形 (431) 4. 隐函数定理 (434) 练习 (438)

§6. 隐函数定理的一些推论 442

1. 反函数定理 (442) 2. 局部地把光滑映射化为典则形式 (446) 3. 函数相关性 (450) 4. 局部地分解微分同胚为最简形式的复合 (451) 5. 莫尔斯引理 (453) 练习 (456)

§7. $\mathbb{R}^n$ 中的曲面和条件极值理论 457

1. $\mathbb{R}^n$ 中的 k 维曲面 (458) 2. 切空间 (462) 3. 条件极值 (466) 练习 (477)

口试试题 481

考试大纲 486

参考文献 489

名词索引 494

中文版修订者的话. 509

第一章　一些通用的数学概念与记号

§1. 逻辑符号

1. 关系与括号　本书的语言，像大多数数学教科书那样，是由普通的语言及一串用以叙述理论的专用符号构成的. 除了这些按照需要而引入的专用符号之外，我们还要利用通用的数学逻辑符号$\neg, \wedge, \vee, \Longrightarrow$、$\Longleftrightarrow$，它们分别表示否定词“非”，系词“与”，“或”，“蕴含”，“等价”[1].

作为例子，我们举出代表三种不同旨趣的意见：

L.“如果采用适合于发现的记号，…… 那么，思考工作就能得到惊人的简化.”(莱布尼茨[2])

P.“数学是把不同实体统一命名的艺术.”(庞加莱[3])

G.“自然界这部巨著是用数学语言写成的.”(伽利略[4])

于是，根据上述记号，有下页开头表的写法.

我们将会看到，只使用形式化记号，回避普通语言，并非总是明智的.

另外，我们发现，由较简单的命题构成复杂命题时，使用了括号，正像写代数式那样，它们起着有关结构层次的作用. 同代数中一样，为了节省括号，可以约定“运

---

[1] 在逻辑学中常使用符号 &，而不用 $\wedge$. 蕴含符号 $\Longrightarrow$ 常写成 $\to$，而等价关系写做 $\longleftrightarrow$ 或 $\leftrightarrow$. 但是为了不更换分析中传统的极限记号 $\to$，在本教程中仍使用上面给出的记号.

[2] 莱布尼茨(Leibniz)(1646—1716) 是杰出的德国学者、哲学家和数学家. 他与牛顿一起享有发现无穷小分析基础的荣誉.

[3] 庞加莱(Poincaré)(1857—1912) 是法国数学家，他的卓越思想革新了数学的许多部门并在数学物理中得到具有重大价值的应用.

[4] 伽利略(Galileo)(1564—1642) 是意大利学者，伟大的自然科学家. 他的工作成为后来关于空间和时间的全部物理概念的基础. 他是现代物理学之父.

写　法	表　示
$L \Rightarrow P$	L 蕴含 P
$L \Leftrightarrow P$	L 与 P 等价
$((L \Rightarrow P) \wedge (\neg P)) \Rightarrow (\neg L)$	若 P 由 L 推出, 而 P 不真, 则 L 不真.
$\neg((L \Leftrightarrow G) \vee (P \Leftrightarrow G))$	G 既不等价于 L, 也不等价于 P.

算次序". 为此目的, 我们对符号规定如下的优先次序:

$$\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow .$$

在这样约定下, 式子 $\neg A \wedge B \vee C \Rightarrow D$ 应解释成 $(((\neg A) \wedge B) \vee C) \Rightarrow D$, 而 $A \vee B \Rightarrow C$ 应解释成 $(A \vee B) \Rightarrow C$, 而不是 $A \vee (B \Rightarrow C)$.

对于表示 A 蕴含 B 或 B 由 A 推出的写法 $A \Rightarrow B$, 我们常常赋予它另一种文字解释:B 是 A 的*必要特征*或*必要条件*, 同样地, A 是 B 的*充分特征*或*充分条件*. 于是关系 $A \Leftrightarrow B$ 可用下面任一种方法去解释:

对于B, A既是必要的又是充分的;

A成立, 当且仅当B成立;

当且仅当B成立,有A成立;

A与B等价.

因此, 写法 $A \Leftrightarrow B$ 表示 A 蕴含 B, 同时 B 蕴含 A.

对式子 $A \wedge B$ 中连接词 "与" 的用法不需作解释了.

然而应注意, 在式子 $A \vee B$ 中的连接词 "或", 不是区分连接词, 也就是说只要命题 A, B 中有一个为真, $A \vee B$ 就正确. 例如, 设 x 是使 $x^2 - 3x + 2 = 0$ 的实数. 这时可以说下列关系成立:

$$(x^2 - 3x + 2 = 0) \Leftrightarrow (x = 1) \vee (x = 2).$$

2. 关于证明的注记　典型的数学论断具有 $A \Rightarrow B$ 这种形式, 这里 A 是前提, B 是结论. 证明这个论断, 就是要建立一串蕴含关系

$$A \Rightarrow C_1 \Rightarrow \cdots \Rightarrow C_n \Rightarrow B,$$

其中每个蕴含关系或为公理, 或为已证明了的断语[①].

在证明中, 我们将使用古典的推证法则: 若 A 真且 $A \Rightarrow B$, 那么 B 也真.

在用反证法证明时, 我们还将使用排中律, 根据它, 不论命题 A 的内容是什么, $A \vee \neg A$(A 或非 A) 总是正确的. 因此, 我们同时还采用 $\neg(\neg A) \Leftrightarrow A$, 即一个命题两次否定等价于原命题.

[①] $A \Rightarrow B \Rightarrow C$ 是 $(A \Rightarrow B) \wedge (B \Rightarrow C)$ 的缩写.

3. 某些专门记号 为方便读者及简化文字叙述, 约定分别用记号 ◀ 及 ▶ 来表示证明的开始和结束.

还约定, 当方便时, 用专门的记号:=(据定义等于) 引进定义, 其中两点放在被定义的对象一边.

例如, 式子

$$\int_a^b f(x)dx := \lim_{\lambda(P)\to 0} \sigma(f, P, \xi)$$

是用右端定义左端. 而右端的含义认为是已知的.

类似地, 对已有定义的式子, 也用这个记号引进简缩记法. 例如, 式子

$$\sum_{i=1}^{n} f(\xi_i)\Delta x_i =: \sigma(f, P, \xi)$$

就是引进记号 $\sigma(f, P, \xi)$ 表示左端的专门和式.

4. 最后的注记 注意, 在这里我们实际上只谈到了记法, 并没有分析逻辑推导形式, 也没涉及诸如真实性、可证明性、可推导性等构成数理逻辑研究对象的深刻问题.

如果我们没有形式逻辑, 怎么建立数学分析呢? 这里使人得到一些安慰的是, 我们知道的, 或说得更恰当些, 我们会的, 总比做形式化时所需要的多些. 这可以用一句格言来说明: 当你要求一只蜈蚣解释它是怎样控制那么多条腿的时候, 它早已学会走路了.

整个科学的经验使我们断定, 昨天认为明显、简单且不能分的东西, 今天可能受到重新审查或把它精确化. 数学分析的许多概念就曾经是这样的 (无疑, 未来也是这样). 它的最重要的一些定理和工具早在 17、18 世纪就已经发现, 然而, 只是在创立了极限理论, 以及这个理论所必需的、逻辑上完全合格的实数理论 (在 19 世纪) 之后, 数学分析才获得了现代形式化的、含义确切的、从而为人们理解的形式.

在第二章中, 我们正是从实数理论的这种水平上开始建造数学分析的整个大厦.

如序言所说, 希望尽快了解微积分学本身的基本概念和有效方法的读者, 可立刻从第三章开始学, 只是在有必要时, 再回到头两章查阅相应的地方.

练 习

用 1 表示命题正确, 用 0 表示命题不正确. 这时, 对命题 $\neg A, A\wedge B, A\vee B, A\Rightarrow B$ 中的每一个, 可建立一个所谓真假表, 它依据命题 A, B 的真实性, 指出这个命题的真实性. 这些表是逻辑运算 $\neg, \wedge, \vee, \Rightarrow$ 的形式定义. 见下页所画的四个表.

1. 验证这些表中所指示的一切是否都与你关于相应的逻辑运算的通常观念相符号.(要特别注意, 若 A 不真, 则 $A\Rightarrow B$ 总是真的.)

2. 试证下面的简单关系成立, 它们在数学的论证中极为重要且有广泛的应用.

¬A

A	0	1
$\neg A$	1	0

A ∧ B

$A \backslash B$	0	1
0	0	0
1	0	1

A ∨ B

$A \backslash B$	0	1
0	0	1
1	1	1

A ⇒ B

$A \backslash B$	0	1
0	1	1
1	0	1

a) $\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$;

b) $\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B$;

c) $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$;

d) $(A \Rightarrow B) \Leftrightarrow \neg A \vee B$;

e) $\neg(A \Rightarrow B) \Leftrightarrow A \wedge \neg B$.

§2. 集与集的初等运算

1. 集合的概念 从 19 世纪末到 20 世纪初, 集合论语言成为最通用的数学语言. 这甚至表现在对数学所下的一种定义中, 它说: 数学就是研究集合上各种结构 (关系) 的科学①.

"所谓集合, 是我们直观感到或意识到的, 由确定的、彼此不同的对象联合成的整体"; 集合论的奠基人乔治 · 康托尔②就是这样描述集合概念的.

康托尔的描述, 当然不能叫做定义, 因为它诉诸于可能比集合概念本身更复杂且从未定义过的概念. 这种描述的目的是把这个概念与其他概念联系起来加以说明.

康托尔的 (或叫做 "朴素的") 集合论的基本前提可归结为:

1° 集合可由任意不同的事物组成;

2° 集合由构成它的事物集聚而唯一确定;

3° 任何性质都定义一个具有该性质的事物的集合.

若 x 是一事物, P 是一性质, $P(x)$ 表示 x 有性质 P. 用 $\{x|P(x)\}$ 表示具有性质 P 的一切事物的类. 组成类或集合的事物, 叫做类或集合的元素.

由元素 $x_1, \cdots, x_n$ 组成的集合用 $\{x_1, \cdots, x_n\}$ 表示. 当不致引起混淆时, 为了书写简单, 我们直接用 a 表示单元素集合 $\{a\}$.

"类","族","集体","组" 等字, 在朴素集合论中作 "集合" 的同义词来使用.

---

①数学的结构. 参看 Bourbaki 的书《数学简史》, 俄译本 Очерки по истории математики, M.: ИА,1963.

②康托尔(G. Cantor)(1845—1918) —— 德国数学家, 无穷集理论的创始人, 在数学中使用集合论语言的鼻祖.

下面的例子说明这些术语的应用:

在词 "я" 中的字母 "a" 的集合;

阿达马的妻子的集合;

十个数码所成的组;

豆科植物族;

地球上沙粒的集合;

平面上与其上二已知点等距离的点的全体;

集合的族;

所有集合的集合.

集合课题的确定性在程度上可能具有的差别, 向我们提示, 集合这个概念, 可不是那么简单和不出麻烦的概念.

事实上, 例如一切集合的集合这个概念, 就会产生矛盾.

◄ 设 M 为一集合, 用 $P(M)$ 表示 "M 是不以自己作为元素的集" 这样一种性质.

考察具有性质 P 的集合的类 $K = \{M|P(M)\}$.

如果 K 是集合, 那么, 或者 $P(K)$ 为真, 或者 $\neg P(K)$ 为真. 然而, 二者择一对于 K 是不可能的. 实际上, $P(K)$ 不成立, 因为由 K 的定义推知 K 包含着 K, 即 $\neg P(K)$ 为真; 另一方面, $\neg P(K)$ 也不可能真, 因为这就表示 K 包含着 K, 而这与 K 的定义, 亦即, 它是不含自身的那样的集的类, 相矛盾.

因此, K 不是集合 ►

这是经典的罗素[①]悖论, 它是朴素集合论所导致的悖论之一.

在近代数理逻辑中, 集合的概念受到了精细的推敲 (我们会看到, 这并不是无根据的). 但是, 我们不去进行这种深人的分析. 我们仅指出, 在现行的公理化理论中, 集合被定义为有一套确定性质的数学对象.

描述这些性质构成了公理. 集合论公理系统的核心是假定了一些规则, 根据这些规则可以从一些集合构成新的集合. 总之, 现行的任何公理体系, 一方面, 要避免朴素理论中众所周知的矛盾, 另一方面, 要保证适应于各种具体的集合, 这些集合来源于数学的不同部门, 首先是数学分析, 当然是按广义理解下的数学分析.

我们对集合的概念, 暂时就给出这些注释, 并转而描述在分析中最常用的集合性质.

希望对集合概念作更详细地了解的读者, 可参看本章 §4 的第二段, 或查阅专门文献.

2. 包含关系 已经说过, 组成集合的事物, 叫做该集合的*元素*. 我们尽量地用大写拉丁字母表示集合, 而用对应的小写字母表示集合的元素.

[①]罗素 (B. Russell)(1872—1970) —— 英国逻辑学家, 哲学家, 社会学者, 社会活动家.

命题 "x 是集合 X 的元素" 用符号简单地记作

$$x \in X(\text{或}X \ni x),$$

它的否命题用符号

$$x \notin X \ (\text{或}X \not\ni x)$$

来记.

在书写有关集合的命题时, 经常运用逻辑运算 $\exists$("存在" 或 "找到") 与 $\forall$("任何的" 或 "对于任何的"), 分别称之为*存在量词*与*全称量词*.

例如写法 $\forall x((x \in A) \Leftrightarrow (x \in B))$ 表示的是, 对于任何事物 x, 关系 $x \in A$ 与 $x \in B$ 是等价的. 因为一个集合完全被它的元素所确定, 所以, 上述命题可以简记为

$$A = B,$$

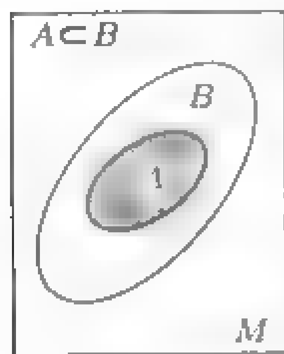


图 1

读做 "A 等于 B", 它表示集合 A 与集合 B 完全一致.

这样, 当二集合由同样的元素构成时, 它们相等.

否定相等就写成 $A \neq B$.

若集合 A 的任何元素都是集合 B 的元素, 就记作 $A \subset B$ 或 $B \supset A$, 并说, 集合 A 是集合 B 的*子集*, 或说 B 包含 A, 或说 B 把 A 包含于自己内. 因此, 集合 A, B 间的关系 $A \subset B$ 叫做*包含关系*(图 1).

于是,

$$(A \subset B) := \forall x((x \in A) \Rightarrow (x \in B)).$$

如果 $A \subset B$ 且 $A \neq B$, 就说包含关系 $A \subset B$ 是*严格的*, 或说 A 是 B 的*真子集*. 利用所引进的定义, 可以推知

$$(A = B) \Leftrightarrow (A \subset B) \wedge (B \subset A).$$

如果 M 是集合, P 是任一性质, 那么, 就能从 M 中分出一个子集

$$\{x \in M | P(x)\},$$

它是由 M 中具有这一性质的一切元素组成的子集.

例如, 显然有

$$M = \{x \in M | x \in M\}.$$

另一方面, 如果 P 是 M 中任何元素都不具有的一个性质, 比如说, $P(x) := (x \neq x)$, 则我们得到集合

$$\varnothing = \{x \in M | x \neq x\},$$

它叫做 M 的*空子集*.

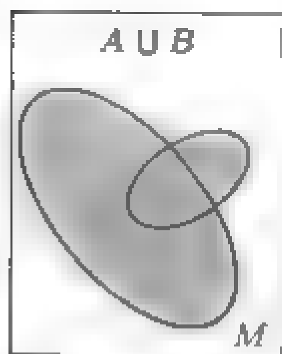


图 2

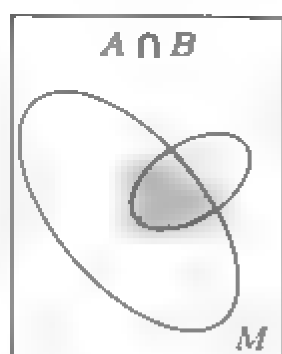


图 3

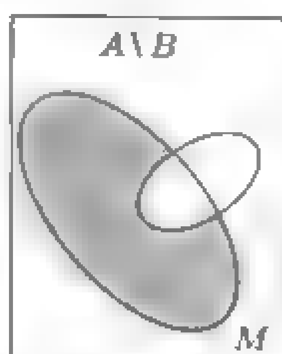


图 4

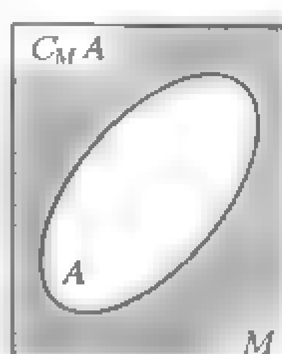


图 5

3. 最简单的集合运算 设 A 与 B 都是集合 M 的子集.

a. 集合

$$A\cup B := \{x\in M|(x\in A)\vee(x\in B)\},$$

由 M 中的那些元素组成, 它至少属于 A, B 中之一 (图 2), 称此集为 A, B 的并集.

b. 集合

$$A\cap B := \{x\in M|(x\in A)\wedge(x\in B)\},$$

由 M 中的那些元素组成, 它同时属于 A 和 B(图 3), 称此集为 A, B 的交集.

c. 集合

$$A\backslash B := \{x\in M|(x\in A)\wedge(x\notin B)\},$$

由 M 中的那些元素组成, 它属于 A 但不属于 B(图 4), 称此集为 A, B 的差集.

集 M 与 M 的子集 A 之差通常叫做 A 在 M 中的补集, 记作 C_MA, 如果由前后文能明显地知道 A 是哪个集中的补集时, 就记作 CA(图 5)

例 为了说明引进的这些概念的相互作用, 我们验证下面的关系式 (称之为德 · 摩根[①] 规则):

$$C_M(A\cup B) = C_MA\cap C_MB, \tag{1}$$

$$C_M(A\cap B) = C_MA\cup C_MB. \tag{2}$$

◀ 例如我们来证明第一个等式:

$$\begin{aligned}(x\in C_M(A\cup B)) &\Rightarrow (x\notin A\cup B)\\ &\Rightarrow ((x\notin A)\wedge(x\notin B))\\ &\Rightarrow (x\in C_MA)\wedge(x\in C_MB)\\ &\Rightarrow (x\in(C_MA\cap C_MB)).\end{aligned}$$

因此, 得到

$$C_M(A\cup B)\subset(C_MA\cap C_MB). \tag{3}$$

[①] 德 · 摩根(De. Morgan)(1806—1871)—— 苏格兰数学家.

另一方面,

$$\begin{aligned}(x \in (C_M A \cap C_M B)) &\Rightarrow ((x \in C_M A) \wedge (x \in C_M B)) \\ &\Rightarrow ((x \notin A) \wedge (x \notin B)) \\ &\Rightarrow (x \notin (A \cup B)) \\ &\Rightarrow (x \in C_M (A \cup B)),\end{aligned}$$

即

$$C_M(A \cup B) \supset (C_M A \cap C_M B). \tag{4}$$

由 (3) 与 (4) 即得到 (1). ►

d. *集合的直积* (*笛卡儿积*). 对于任意两个集合 A, B, 可构造出新的集合 —— 对 $\{A, B\} = \{B, A\}$, 它的元素只有集合 A 和 B. 如果 $A \neq B$, 它有两个元素, 而当 $A = B$ 时, 它只有一个元素.

这个新集合叫做 A, B 的*无序对*, 以区别于*序对*(A, B). 在序对 (A, B) 中, 元素 A, B 被赋予一种附加的特征, 根据它能分辨对 $\{A, B\}$ 中的第一个元素和第二个元素. 按定义, 序对等式

$$(A, B) = (C, D),$$

表示 $A = C$ 且 $B = D$. 特别地, 如果 $A \neq B$, 那么 $(A, B) \neq (B, A)$.

现设 X, Y 是任意两个集. 按定义, 由一切序对 (x, y)(其第一项是 X 中的元素, 而第 2 项是 Y 中的元素) 所构成的集合

$$X \times Y := \{(x, y) | (x \in X) \wedge (y \in Y)\},$$

叫做集合 X, Y(按这样的次序!) 的*直积*或*笛卡儿积*.

由直积的定义及上面关于序对所做的注释, 一般来说, $X \times Y \neq Y \times X$. 只有当 $X = Y$ 时, 等式才成立. 这时我们把 $X \times X$ 简记作 X^2.

直积又叫做笛卡儿积, 以纪念笛卡儿[①], 他曾与费马[②]各自独立地通过坐标系引进了几何的分析语言. 众所周知的平面笛卡儿坐标系, 恰好把平面变成了两个数轴的直积. 笛卡儿积与它的因子的次序有关这一点, 在这个熟知的情形是非常显然的. 例如, 序对 (0,1) 和 (1,0) 对应于平面上两个不同的点.

设序对 $z = (x_1, x_2)$ 是集合 X_1, X_2 的直积$X_1 \times X_2$ 中的元素, x_1 叫做序对 z 的*第一射影*, 记作 $\mathrm{pr}_1 z$; 而 x_2 叫做序对 z 的*第二射影*, 记作 $\mathrm{pr}_2 z$.

---

①笛卡儿(Descartes)(1596—1650)—— 著名的法国哲学家, 数学家与物理学家. 他在科学思想方法和认识论方面作出了重大贡献.

②费马(Fermat)(1601—1665)—— 卓越的法国数学家, 职业法学家. 费马在分析, 解析几何, 概率论, 数论等一系列现代数学领域中, 都居于创始的地位

练 习

在问题 1,2,3 中用 A, B, C 表示某集 M 的子集.

1. 验证下面诸关系式:
   a) $(A \subset C) \wedge (B \subset C) \Leftrightarrow ((A \cup B) \subset C)$;
   b) $(C \subset A) \wedge (C \subset B) \Leftrightarrow (C \subset (A \cap B))$;
   c) $C_M(C_M A) = A$;
   d) $(A \subset C_M B) \Leftrightarrow (B \subset C_M A)$;
   e) $(A \subset B) \Leftrightarrow (C_M A \supset C_M B)$.

2. 试证:
   a) $A \cup (B \cup C) = (A \cup B) \cup C =: A \cup B \cup C$;
   b) $A \cap (B \cap C) = (A \cap B) \cap C =: A \cap B \cap C$;
   c) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$;
   d) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

3. 验证并与交运算的对偶关系:
   a) $C_M(A \cup B) = C_M A \cap C_M B$;
   b) $C_M(A \cap B) = C_M A \cup C_M B$.

4. 对下列集合的笛卡儿积作出几何解释:
   a) 二线段 (矩形);
   b) 二直线 (平面);
   c) 直线与圆周 (圆柱面);
   d) 直线与圆面 (圆柱体);
   e) 二圆周 (圆环面);
   f) 圆周与圆面 (实圆环).

5. 集合 $\Delta = \{(x_1, x_2) \in X^2 | x_1 = x_2\}$ 叫做集合 X 的笛卡儿正方形 X^2 的对角线.

   对第 4 题的 a),b),e) 所得之集的对角线做出几何解释.

6. 试证:
   a) $(X \times Y = \varnothing) \Leftrightarrow (X = \varnothing) \vee (Y = \varnothing)$,

   而如果 $X \times Y \neq \varnothing$, 则

   b) $(A \times B \subset X \times Y) \Leftrightarrow (A \subset X) \wedge (B \subset Y)$.
   c) $(X \times Y) \cup (Z \times Y) = (X \cup Z) \times Y$.
   d) $(X \times Y) \cap (X' \times Y') = (X \cap X') \times (Y \cap Y')$.

   这里 $\varnothing$ 表空集, 即不含任何元素的集合.

7. 将问题 3 中的关系与 §1 练习 2 中的关系 a),b) 比较, 建立对于命题的逻辑运算$\neg$、$\wedge$、$\vee$ 与对于集合的运算$C, \cap, \cup$ 之间的关系.

§3. 函数

1. 函数 (映射) 的概念 现在来描述函数关系的概念, 它不仅对于数学是基本的.

设有二集合 X 与 Y.

如果按照某种规律 f, 对于每个元素 $x \in X$, 有一元素 $y \in Y$ 与之对应, 就说有一个定义在 X 上而在 Y 中取值的函数.

这时, 称集合 X 为函数的定义域; 它的一般元素 x 称为函数的变元或自变量; 与变元 x 的具体值 $x_0 \in X$ 相对应的元素 $y_0 \in Y$, 叫做在元素 x_0 上的函数值, 或叫做当变元 $x = x_0$ 时的函数值, 并以 $f(x_0)$ 表示. 当变元变化时, 一般来说, 值 $y = f(x) \in Y$ 随 x 的值而变. 因此, 常称量 $y = f(x)$ 为函数的因变量.

函数在集合 X 的一切元素上取得的所有函数值的集合

$$f(X) := \{y \in Y | \exists x((x \in X) \wedge (y = f(x)))\}$$

叫做函数的值集或函数的值域.

由于集合 X, Y 的性质不同, 术语 "函数" 在不同的数学部门中各有其合用的同义词:映射,变换, 射,算子,泛函. 其中映射是最通用的, 我们也经常使用这个名词.

常用记号

$$f : X \to Y, X \xrightarrow{f} Y.$$

来记函数 (映射).

如果函数的定义域和函数值域从上下文看很明显时, 也用

$$x \mapsto f(x) \quad \text{或} \quad y = f(x)$$

这种简单记法, 比较常见的是, 就只用一个字母 f 来表示函数.

如果两个函数 f_1, f_2 有相同的定义域 X, 且对每个 $x \in X$, 这两个函数的函数值 $f_1(x), f_2(x)$ 一致, 就认为两个函数 f_1 与 f_2 一致或相等. 这时记作 $f_1 = f_2$.

如果 $A \subset X$, 而 $f : X \to Y$ 是某个函数, 就用 $f|A$ 或 $f|_A$ 来表示函数 $\varphi : A \to Y$, φ 在集合 A 上与 f 一致. 更确切地说, 就是当 $x \in A$ 时, $f|_A(x) := \varphi(x)$. 称函数 $f|_A$ 为函数 f 在集合 A 上的缩小或限制, 而函数 $f : X \to Y$ 相对于函数 $\varphi = f|_A : A \to Y$ 来说, 就叫做函数 φ 在集合 X 上的扩张或延拓.

我们看到, 有时要考察定义在某个集合 X 的子集 A 上的函数 $\varphi : A \to Y$, 并且值域 $\varphi(A)$ 也可能是与 Y 不一致的 Y 的一个子集. 因此, 对于包含定义域的任意集合 X, 有时就用术语 "函数的出发域"称呼它, 而把包含函数值域的任意集合 Y, 叫做 "函数的到达域".

这样, 给出一个函数 (映射)要求指明一个三元组 (X, f, Y), 其中

X—— 函数的被映射集或定义域;

Y—— 函数所映入的集合或函数的到达域;

f—— 规律, 按照它, 对每个 $x \in X$, 有确定的元素 $y \in Y$ 与之对应.

我们看到, 在这里 X 与 Y 是不对称的, 这反映着映射是从 X 到 Y 内的.

现在看看函数的几个例子.

例 1 公式 $l = 2\pi r$ 与 $V = \frac{4}{3}\pi r^3$ 建立了圆周长 l 及球体积 V 对半径 r 的函数依赖关系. 每个公式按其意义各自给出定义在正实数集 $\mathbb{R}_+$ 上, 且在同一集 $\mathbb{R}_+$ 中取值的函数 $f : \mathbb{R}_+ \to \mathbb{R}_+$.

例 2 设 X 是惯性坐标系的集合, 而 $c : X \to \mathbb{R}$ 是这样一个函数: 对于每个惯性坐标系 $x \in X$, 对应一个相对这个惯性坐标系测量出来的真空中的光速 $c(x)$. 函数 $c : X \to \mathbb{R}$ 是常数函数, 即对于任何 $x \in X$, 它有同样的值 c (这是一个基本的实验事实.)

例 3 由公式

$$\begin{cases} x' = x - vt, \\ t' = t, \end{cases}$$

给出到自身的映射 $G : \mathbb{R}^2 \to \mathbb{R}^2$ ($\mathbb{R}^2$ 是时间轴 $\mathbb{R}_t$ 与空间轴 $\mathbb{R}_x$ 的直积: $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \mathbb{R}_t \times \mathbb{R}_x$) 是经典的伽利略变换, 它把一个惯性坐标系 (x, t) 变成另一个惯性坐标系 (x', t'), 后者相对前者的运动速度是 v.

由

$$\begin{cases} x' = \dfrac{x - vt}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}, \\ t' = \dfrac{t - \left(\frac{v}{c^2}\right) x}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}, \end{cases}$$

给出的变换 $L : \mathbb{R}^2 \to \mathbb{R}^2$ 也是为了同样的目的. 这是著名的 (一维)洛伦兹变换 ①, 它在狭义相对论中起着基本的作用; 其中 c 是光速.

例 4 射影 $\mathrm{pr}_1 : X_1 \times X_2 \to X_1$ 由对应 $X_1 \times X_2 \ni (x_1, x_2) \overset{\mathrm{pr}_1}{\longmapsto} x_1 \in X_1$ 给出. 显然, 它是一个函数. 用类似的方法可定义第二个射影 $\mathrm{pr}_2 : X_1 \times X_2 \to X_2$.

例 5 设 $\mathcal{P}(M)$ 是集合 M 的一切子集的集合. 对于每个 $A \in \mathcal{P}(M)$, 令 A 在 M 中的补集 $C_M(A) \in \mathcal{P}(M)$ 与之对应, 于是得到 $\mathcal{P}(M)$ 到自身的映射 $C_M : \mathcal{P}(M) \to \mathcal{P}(M)$.

---

①洛伦兹(G. A. Lorentz)(1853—1928)——荷兰物理学家. 所说的变换是他在 1904 年得到的, 而在 1905 年爱因斯坦表述的狭义相对论中, 这一变换得到了非常重要的应用.

例 6 设 $E \subset M$. 在 M 上用条件

$$(\chi_E(x) = 1, \text{当} x \in E) \wedge (\chi_E(x) = 0, \text{当} x \in C_M(E))$$

定义的实值函数 $\chi_E : M \to \mathbb{R}$, 叫做集合 E 的*特征函数*.

例 7 设 $M(X;Y)$ 是 X 到 Y 中的映射的集合, 而 x_0 为 X 中一个固定的元素. 令每个函数 $f \in M(X;Y)$ 对应于它在 x_0 的值 $f(x_0) \in Y$. 这就确定了一个函数

$$F : M(X;Y) \to Y.$$

特别当 $Y = \mathbb{R}$, 即 Y 为实数集时, 函数 $F : M(X;\mathbb{R}) \to \mathbb{R}$ 把每个函数 $f : X \to \mathbb{R}$ 对应到一个数. $F(f) = f(x_0)$. 这样一来, F 是定义在函数上的函数. 为了方便起见, 把这样的函数叫做*泛函*.

例 8 设 Γ 为位于曲面上 (例如在地球上) 且联接两个固定点的曲线的集合. 可以令 Γ 中每条曲线γ 对应于它的长. 这样由 "曲线的长"就得到一个函数 $F : \Gamma \to \mathbb{R}$, 为了求曲面上二已知点间的最短线或所谓的*测地线*, 常需要研究这种函数.

例 9 考察定义在整个实轴 $\mathbb{R}$ 上的所有实值函数的集合 $M(\mathbb{R};\mathbb{R})$. 固定一个数 $a \in \mathbb{R}$, 对每个函数 $f \in M(\mathbb{R};\mathbb{R})$, 使函数 $f_a \in M(\mathbb{R};\mathbb{R})$ 与它对应, 它们间的联系是 $f_a(x) = f(x+a)$. 函数 $f_a(x)$ 通常叫做将函数 $f(x)$*平移*一个数 a. 这样所生成的映射 $A : M(\mathbb{R};\mathbb{R}) \to M(\mathbb{R};\mathbb{R})$ 叫做*平移算子*. 因此, 算子 A 是在函数上定义的, 而它的值也是函数: $f_a = A(f)$.

如果我们到处都看不到实际的算子, 那么上面考察的这些例子就可能显得好像是人造的. 实际上, 每架收音机就是一个算子 $f \xrightarrow{F} \hat{f}$, 它把电磁信号 f 转变为音频信号 $\hat{f}$; 我们的每个感官, 都是具有它自己的定义域和值域的一个算子 (转换器).

例 10 空间质点的位置是由叫做它的空间坐标的有序三元数组 (x,y,z) 来确定的. 一切这样的有序三元数组构成的集合可理解为三个数轴 $\mathbb{R}$ 的直积 $\mathbb{R}\times\mathbb{R}\times\mathbb{R} = \mathbb{R}^3$.

运动的质点在每一瞬时 t, 位于空间 $\mathbb{R}^3$ 的坐标为 $(x(t), y(t), z(t))$ 的一点. 这样一来, 质点的运动可以解释为映射 $\gamma : \mathbb{R} \to \mathbb{R}^3$, 这里 $\mathbb{R}$ 是时间轴, 而 $\mathbb{R}^3$ 是三维空间.

如果一个系统是由 n 个质点组成, 那么它的构形由各质点的位置所确定, 即由 $3n$ 个数的序组

$$(x_1, y_1, z_1; x_2, y_2, z_2; \cdots; x_n, y_n, z_n)$$

确定, 所有这些序组的集合, 叫做 *n质点系的构形空间*. 因此, n 质点系的构形空间可以解释为 n 个 $\mathbb{R}^3$ 空间的直积

$$\mathbb{R}^3 \times \mathbb{R}^3 \times \cdots \times \mathbb{R}^3 = \mathbb{R}^{3n}.$$

由时间轴到质点系的构形空间的映射 $\gamma : \mathbb{R} \to \mathbb{R}^{3n}$ 对应于 n 质点系的运动.

例 11 力学系统的势能 U 与其质点的分布有关, 即与它的构形有关. 设 Q 是力学系统的实际可能的构形的集合. 这是系统的构形空间的某个子集. 每个位置 $q \in Q$ 对应着系统势能的一个值 $U(q)$, 因此, 势能是定义在构形空间的子集 Q 上, 而在实数域 $\mathbb{R}$ 中取值的一个*函数*$U : Q \to \mathbb{R}$.

例 12 n 质点系的动能K 与它们的速度有关. 因此, 质点系的全部机械能 $E = K + U$, 即动能与势能之和, 它既与质点系的构形有关, 又与诸质点的速度组 v 有关. 像质点系在空间的构形 q 一样, 由 n 个三维向量构成的速度组 v, 也能用 $3n$ 个数的序组给出. 与我们的质点组的状态对应的序对 (q, v) 构成直积空间 $\mathbb{R}^{3n} \times \mathbb{R}^{3n} = \mathbb{R}^{6n}$ 中的一个子集 Φ, 叫做 *n质点系的相空间*(它与 $\mathbb{R}^{3n}$ 中的构形空间是不同的).

因此, 系统的总能量是定义在相空间 $\mathbb{R}^{6n}$ 的子集 Φ 上, 而在实数域 $\mathbb{R}$ 中取值的*函数* $E : \Phi \to \mathbb{R}$.

特别地, 当系统是封闭系统, 即没有任何外力作用于它时, 根据能量守恒定律, 函数 E 在系统状态集 Φ 的所有点上有同样的值 $E_0 \in \mathbb{R}$.

2. 映射的简单分类 当把函数 $f : X \to Y$ 叫做映射时, 它在元素 $x \in X$ 上取的值 $f(x) \in Y$, 通常称为*元素 x 的像*.

设 $A \subset X$, 则把 A 中各点 x 的像 (在 Y 中) 所组成的集

$$f(A) := \{y \in Y | \exists x((x \in A) \wedge (y \in f(x))\}$$

叫做 A 在映射 $f : X \to Y$ 下的像.

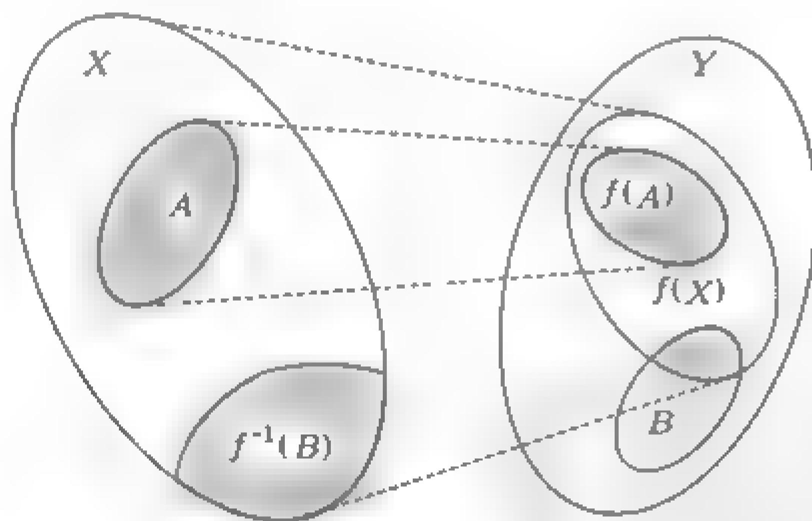


图 6

设集 $B \subset Y$, 则把 X 中其像属于 B 的那些元素的集

$$f^{-1}(B) := \{x \in X | f(x) \in B\}$$

叫做 B 的*原像*或*全原像*(图 6)

关于映射 $f : X \to Y$:

如果 $f(X)=Y$, 就说 f 是满射(或 X 到 Y 上的映射);

如果对 X 中的任何元素 x_1, x_2, 有

$$(f(x_1)=f(x_2)) \Rightarrow (x_1=x_2),$$

即不同元素有不同的像, 就说 f 是单射 (或嵌入, 内射);

如果 f 既是满射又是单射, 就说 f 是双射(或一一映射又叫双方单值映射).

如果映射 $f: X \to Y$ 是双射, 即是集合 X 与集合 Y 间的双方单值对应, 那么, 自然可以做出一个映射

$$f^{-1}: Y \to X,$$

它用以下方法定义: 如果 $f(x)=y$, 则 $f^{-1}(y)=x$, 亦即, 与$y \in Y$ 对应的是那样的元素 $x \in X$, 它在 f 下的像是 x. 根据 f 的满射性, 这样的 $x \in X$ 存在; 而由 f 的内射性, 它又是唯一的. 因此, 映射 f^{-1} 完全确定. 这个映射叫做原映射 f 的逆映射.

由逆映射的做法看出, $f^{-1}: Y \to X$ 本身也是双射, 并且它的逆映射 $(f^{-1})^{-1}: X \to Y$ 与 $f: X \to Y$ 一致.

因此, 两个映射是逆映射的性质是相互的, 即如果 f^{-1} 是 f 的逆映射, 反过来, f 也是 f^{-1} 的逆映射.

我们注意到, 集 $B \subset Y$ 的原像的符号 $f^{-1}(B)$ 与反函数的记号 f^{-1} 有联系, 然而, 应该看到, 集合的原像是对任何映射 $f: X \to Y$ 都有定义的, 即使它不是双射, 从而没有逆.

3. 函数的复合与互逆映射 函数的复合运算, 一方面是产生新函数的丰富源泉, 另一方面又是将复杂函数分解成比较简单的函数的一种方法.

若有二映射 $f: X \to Y$ 与 $g: Y \to Z$, 且 g 定义在 f 的值域上, 则可用公式

$$(g \circ f)(x) := g(f(x))$$

确定 X 上的新映射

$$g \circ f: X \to Z.$$

所建立的这个映射 $g \circ f$ 叫做映射 f 与映射 g(按这种顺序!) 的复合映射.

图 7 具体说明了映射 f 与 g 复合的结构.

无论在几何中讨论了平面运动或空间运动的复合时, 还是在代数中研究由最简单的初等函数复合得到的复杂函数时, 我们都不止一次地遇到过映射的复合.

复合运算有时要连续施行若干次, 因而, 注意到它们是结合的, 即

$$h \circ (g \circ f) = (h \circ g) \circ f$$

是有好处的.

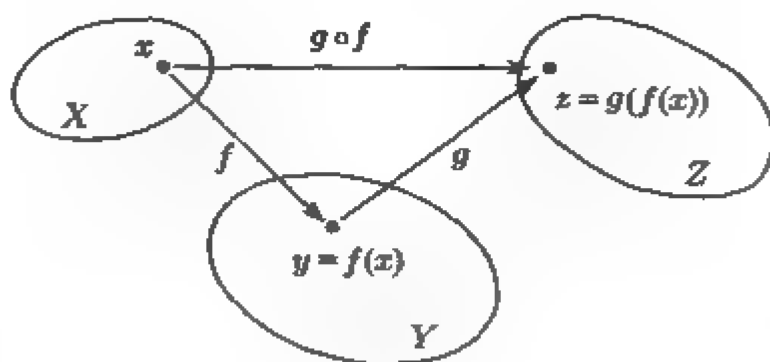


图 7

◀ 实际上,

$$\begin{aligned}(h\circ(g\circ f))(x) &= h((g\circ f)(x)) = h(g(f(x))) \\ &= (h\circ g)(f(x)) = ((h\circ g)\circ f)(x).\end{aligned}$$ ▶

因此, 正像几个数的加法与乘法那样, 可以去掉表明运算次序的括号.

如果复合 $f_n\circ\cdots\circ f_1$ 中所有的项都等于 f, 那么就把它简短地表成 f^n.

大家都知道, 例如, 正数 a 的平方根, 可以按照公式

$$x_{n+1}=\frac{1}{2}\left(x_n+\frac{a}{x_n}\right),$$

用逐次逼近法从任意的第一近似值 $x_0>0$ 开始进行计算. 这恰好就是在计算 $f^n(x_0)$, 这里 $f(x)=\frac{1}{2}\left(x+\frac{a}{x}\right)$. 这种把前一步得到的函数值, 作为下一步的自变量值的计算方法叫做迭代法. 在数学中, 迭代法用处很广.

我们还注意到, 即使两种复合 $g\circ f$ 与 $f\circ g$ 都有定义, 一般来说, 它们也是不相等的:

$$g\circ f\neq f\circ g.$$

实际上, 例如取二元集 $\{a,b\}$, 及映射 $f:\{a,b\}\to a, g:\{a,b\}\to b$. 这时, 显然 $g\circ f:\{a,b\}\to b$, 而 $f\circ g:\{a,b\}\to a$.

如果映射 $f:X\to X$ 把 X 的每个元映成自身, 即 $x\xrightarrow{f}x$, 那么就把 f 记作 e_X, 并称之为集 X 的恒等映射.

引理

$$(g\circ f=e_X)\Rightarrow(g\text{是满射})\wedge(f\text{是单射}).$$

◀ 实际上, 如果 $f:X\to Y; g:Y\to X$, 且 $g\circ f=e_X:X\to X$, 那么

$$X=e_X(X)=(g\circ f)(X)=g(f(X))\subset g(Y),$$

也就是说, g 是满射.

此外, 如果 $x_1 \in X, x_2 \in X$, 则

$$(x_1 \neq x_2) \Rightarrow (e_X(x_1) \neq e_X(x_2)) \Rightarrow ((g \circ f)(x_1) \neq (g \circ f)(x_2)) \\ \Rightarrow ((g(f(x_1)) \neq g(f(x_2))) \Rightarrow (f(x_1) \neq f(x_2)),$$

因此, f 是单射. ►

互逆的映射可以用映射的复合来描述.

命题 *映射 $f: X \to Y, g: Y \to X$ 是互逆的双射当且仅当 $g \circ f = e_X$ 且 $f \circ g = e_Y$.*

◄ 据引理, 条件 $g \circ f = e_X$ 与 $f \circ g = e_Y$ 同时成立就保证了映射 f, g 的满射性与单射性, 也就是说, 它们都是双射.

这两个条件又说明 $y = f(x)$ 当且仅当 $x = g(y)$ ►

以上我们用明确构造的方式建立了逆映射. 已证明的这个命题, 使我们能给互逆的映射一个虽不太直观, 然而更对称的定义: 映射 f, g 互为逆映射, 假如它们满足两个条件:

$$g \circ f = e_X, \quad f \circ g = e_Y$$

(参看本节末的习题 6).

4. 作为关系的函数. 函数的图像 最后, 我们再回到函数的概念上来. 我们看到, 它的演变是长期而相当复杂的.

"函数" 这个词首见出现在 1673—1692 年莱布尼茨的著作中 (诚然, 是在较狭窄的意义下). 直到 1698 年, J. 伯努利 (即大伯努利)[①]和莱布尼茨在书信来往中, 才在比较接近近代的意义下使用了这一术语.

几乎与上节开始时引入的函数概念完全一致的描述, 在欧拉那里已经遇到 (18 世纪中叶). 至 19 世纪初, 已经出现在译成俄文的拉克鲁阿[②]的数学教科书中. 罗巴切夫斯基[③]是这一函数概念的积极拥护者. 罗巴切夫斯基还指出,"从一般理论的观点看, 只有把彼此联系的数理解成是一起给定的, 才存在依赖关系" [④], 而这正是我们现在要讲的精确定义函数概念的思想.

在本节开始对函数概念的描述是很有生气的, 并反映了事物的本质. 然而从现代观点来看, 还不能说是一个定义, 因为它利用了与函数等价的概念: 对应. 为使读

[①] J. 伯努利 (J. Bernoulli)(1667—1748) —— 著名的瑞士 Bernoulli 学者家族的早期代表者之一, 他是分析学家, 几何学家, 力学家, 他创立了变分学, 对微分积分计算给出了第一个系统的阐述.

[②] 拉克鲁阿(Lacroix)(1765—1843) — 法国数学家与教师 (师范与综合技术学校教授, 巴黎科学院院士).

[③] 罗巴切夫斯基(Н И Лобачевский)(1792—1856) —— 伟大的俄罗斯学者, 他与伟大的德国自然科学家高斯 (K. Gauss 1777—1855) 以及卓越的匈牙利数学家鲍耶 (J Bolyai, 1802—1860) 一起, 享有以他的名字命名的非欧几何的发现者的荣誉.

[④] 罗巴切夫斯基全集 — М.– Л.: Гостехиздат, 1951 年版第 V 卷第 44 页.

者了解, 在这里, 我们将介绍怎样用集合论语言给出函数定义.(有趣的是, 我们下边介绍的 "关系" 的概念, 莱布尼茨也有, 它是先于函数概念的.)

a. 关系 由一些序对(x,y) 组成的任一集, 叫做一个关系$\mathcal{R}$.

构成 $\mathcal{R}$ 的所有序对的第一个元素组成的集 X 叫做关系 $\mathcal{R}$ 的定义域, 而这些序对的第二个元素组成的集 Y 叫做关系 $\mathcal{R}$ 的值域.

因此, 可把关系 $\mathcal{R}$ 解释成直积 $X \times Y$ 的子集 $\mathcal{R}$. 如果 $X \subset X'$, 且 $Y \subset Y'$, 显然

$$\mathcal{R} \subset X \times Y \subset X' \times Y'.$$

所以同一个关系可作为不同集的子集给出来.

含有某关系的定义域的集, 叫做这个关系的出发域, 而含有其值域的集叫做它的到达域.

常常把 $(x,y) \in \mathcal{R}$ 写成 $x\mathcal{R}y$, 并说 x 与 y 用关系 $\mathcal{R}$ 联系着.

如果 $\mathcal{R} \subset X^2$, 就说关系 $\mathcal{R}$ 在 X 上给定.

现在看几个例子.

例 13 对角线

$$\Delta = \{(a,b) \in X^2 | a = b\}$$

是 X^2 的一个子集, 它给出集 X 中元素的相等关系. 事实上, $a\Delta b$ 表示 $(a,b) \in \Delta$, 即 $a = b$.

例 14 设 X 是平面上的直线构成的集.

两条直线 $a \in X, b \in X$ 被认为由关系 $\mathcal{R}$ 联系着, 写成 $a\mathcal{R}b$, 如果直线 b 平行于直线 a. 显然, 这就从 X^2 中分出了满足 $a\mathcal{R}b$ 的对 (a,b) 的集合 $\mathcal{R}$, 由几何教科书得知, 直线间的平行关系具有下列性质:

$a\mathcal{R}a$ (反身性);

$(a\mathcal{R}b) \Rightarrow (b\mathcal{R}a)$ (对称性);

$(a\mathcal{R}b) \wedge (b\mathcal{R}c) \Rightarrow (a\mathcal{R}c)$ (传递性).

具有上面三条性质, 即反身性[①], 对称性与传递性的任何关系 $\mathcal{R}$, 都叫等价关系. 等价关系用专用的符号 $\sim$ 表示, 在这种情况下, 这个符号代替了表示关系的字母 $\mathcal{R}$. 这样, 在等价关系的情形, 我们以 $a \sim b$ 代替 $a\mathcal{R}b$, 并说a 与 b 等价.

例 15 设 M 是一个集合, 而 $X = \mathcal{P}(M)$ 是它的一切子集的全体. 对于集 $X = \mathcal{P}(M)$ 的任意两个元素 a 与 b, 即 M 的任意两个子集 a 与 b, 下面三种可能总有一种被满足:a 包含在 b 中; b 包含在 a 中; a 不是 b 的子集且 b 也不是 a 的子集.

---

①为完整起见, 我们指出, 关系 $\mathcal{R}$ 称做反身的, 假如它的定义域与值域一致, 并且对于 $\mathcal{R}$ 的定义域中的每个元素 a, 有 $a\mathcal{R}a$ 成立.

我们把 X 的子集间的包含关系作为 X^2 中的关系 $\mathcal{R}$, 即按定义令

$$a\mathcal{R}b := (a \subset b).$$

这个关系显然具有下述性质:

$a\mathcal{R}a$ (反身性);

$(a\mathcal{R}b) \wedge (b\mathcal{R}c) \Rightarrow (a\mathcal{R}c)$ (传递性);

$(a\mathcal{R}b) \wedge (b\mathcal{R}a) \Rightarrow a\Delta b$, 即 $a = b$ (反对称性).

一个集 X 的元素对之间的关系, 如果具有上面所指出的三条性质, 习惯上称它是集 X 上的一个*偏序关系*. 我们经常用记法 $a \preccurlyeq b$ 来表示偏序关系以代替记法 $a\mathcal{R}b$, 并说 *b在 a 之后*.

如果除了确定偏序关系的三条性质之外, 还满足条件:

$$\forall a \forall b((a\mathcal{R}b) \vee (b\mathcal{R}a)),$$

即集 X 中的任二元素均能比较, 就把关系 $\mathcal{R}$ 叫做*序关系*, 而把定义了序关系的集 X 叫做*线性序集*.

这一名称的产生, 与数轴的直观形象有关, 在实数轴上, 对于任何一对实数都能讨论 $\leqslant$ 关系.

b. 函数与函数的图像 如果

$$(x\mathcal{R}y_1) \wedge (x\mathcal{R}y_2) \Rightarrow (y_1 = y_2)$$

就说关系 $\mathcal{R}$ 是一个函数关系.

函数关系叫做*函数*.

特别地, 设 X 与 Y 是两个 (不一定不同的) 集合, $\mathcal{R} \subset X \times Y$ 是定义在 X 上的, 集合 X 的元素 x 与集合 Y 的元素 y 间的一个关系. 如果对于每个 $x \in X$, 恰好只有一个 $y \in Y$ 使得 x 与 y 满足上面的关系, 即 $x\mathcal{R}y$, 那么, $\mathcal{R}$ 就是一个*函数关系*.

这样的函数关系 $\mathcal{R} \subset (X \times Y)$ 是*X 到 Y 中的映射*, 或 *X 到 Y 中的函数*.

我们常用符号 f 来记函数. 如果 f 是函数, 我们将像从前一样书写 $y = f(x)$ 或 $x \overset{f}{\longmapsto} y$ 而不用 xfy, 并把 $y = f(x)$ 叫做函数 f 在元素 x 上的值, 或元素 x 在映射 f 下的像.

对照最初描述函数概念的说法, 我们看到, 按照 "规律"f 把元素 $x \in X$"对应" 于元素 $y \in Y$ 就是对每个 $x \in X$, 指出了唯一的 $y \in Y$, 使 xfy, 即 $(x, y) \in f \subset X \times Y$.

设 Γ 是直积 $X \times Y$ 的子集, 它由一切形如 $(x, f(x))$ 的元素组成, 因而

$$\Gamma := \{(x, y) \in X \times Y | y = f(x)\}.$$

我们称这个子集 Γ 是在原来意义下函数 $f : X \to Y$ 的图像.

在我们把函数作为子集 $f\subset X\times Y$ 这种新的叙述下，函数与它的图像当然已经没有区别了.

我们原则上指出了，形式地建立函数的形式化集合论定义的可能性，实质上是归结为函数与其图像等同. 然而我们并不打算今后只限于用这种形式给出函数. 为了给出函数关系，有时用解析式方便，有时用函数值表方便，也有时用运算过程（算法）的语言描述更方便，根据这种描述，对于给定的 $x\in X$，可以找出相应的 $y\in Y$. 无论用上面哪种方式给出函数，利用其图像给出函数，即作函数图像的问题都是有意义的. 用很精致的图像来给出数值函数的益处是，它能直观地表现出函数依赖关系的基本的定性的特征. 函数的图像（诺谟图），也可以用于计算，但通常只是在精度要求不高时才用到. 精确计算时要用函数值表，而更常见的是借助于算法在计算机上进行.

练　习

1. 用下列方式来定义两个关系 $\mathcal{R}_1,\mathcal{R}_2$ 的复合 $\mathcal{R}_2\circ\mathcal{R}_1$：

$$\mathcal{R}_2\circ\mathcal{R}_1 := \{(x,z)|\exists y(x\mathcal{R}_1y)\wedge(y\mathcal{R}_2z)\}.$$

特别地，当 $\mathcal{R}_1\subset X\times Y,\mathcal{R}_2\subset Y\times Z$ 时，有 $\mathcal{R}=\mathcal{R}_2\circ\mathcal{R}_1\subset X\times Z$，并且

$$x\mathcal{R}z := \exists y((y\in Y)\wedge(x\mathcal{R}_1y)\wedge(y\mathcal{R}_2z)).$$

a) 设 Δ_X 是集合 X^2 的对角线，Δ_Y 是集合 Y^2 的对角线. 试证，如果关系 $\mathcal{R}_1\subset X\times Y$ 与 $\mathcal{R}_2\subset Y\times X$ 满足 $(\mathcal{R}_1\circ\mathcal{R}_2=\Delta_X)\wedge(\mathcal{R}_2\circ\mathcal{R}_1=\Delta_Y)$，那么，它们都是函数，并且给出了集合 X,Y 间的互逆映射.

b) 设 $\mathcal{R}\subset X^2$. 试证：关系 $\mathcal{R}$ 的传递性条件等价于 $\mathcal{R}\circ\mathcal{R}\subset\mathcal{R}$.

c) 如果 $(y\mathcal{R}'x)\Leftrightarrow(x\mathcal{R}y)$，就说关系 $\mathcal{R}'\subset Y\times X$ 是 $\mathcal{R}\subset X\times Y$ 的*转置关系*.
试证，关系 $\mathcal{R}\subset X^2$ 的反对称性等价于条件 $\mathcal{R}\cap\mathcal{R}'\subset\Delta_X$.

d) 验证，集合 X 的任意二元素由关系 $\mathcal{R}\subset X^2$ 相联系（按这种或那种次序），当且仅当 $\mathcal{R}\cup\mathcal{R}'=X^2$.

2. 设 $f:X\to Y$ 是映射. 元素 $y\in Y$ 的原像 $f^{-1}(y)\subset X$ 叫做 y 上的*层*.

a) 指出映射

$$\mathrm{pr}_1:X_1\times X_2\to X_1,\mathrm{pr}_2:X_1\times X_2\to X_2$$

的层是什么.

b) 设 $x_1\in X,x_2\in X$，如果 $f(x_1)=f(x_2)$，即如果 x_1 与 x_2 位于同一个层中，就认为 $x_1\in X$ 与 $x_2\in X$ 由关系 $\mathcal{R}\subset X^2$ 相联系，并且写成 $x_1\mathcal{R}x_2$.
验证 $\mathcal{R}$ 是等价关系

c) 试证映射 $f:X\to Y$ 的层互不相交，而所有层的并是整个集 X.

d) 验证：对集合的元素间的任何等价关系，可将该集表成互不相交的等价元素类的并的形式.

3. 设 $f: X \to Y$ 是 X 到 Y 中的映射. 试证, 如果 A 与 B 都是 X 的子集, 那么

a) $(A \subset B) \Rightarrow (f(A) \subset f(B)) \nRightarrow (A \subset B)$,

b) $(A \neq \varnothing) \Rightarrow (f(A) \neq \varnothing)$,

c) $f(A \cap B) \subset f(A) \cap f(B)$,

d) $f(A \cup B) = f(A) \cup f(B)$;

如果 A' 与 B' 都是 Y 的子集, 那么

e) $(A' \subset B') \Rightarrow f^{-1}(A') \subset f^{-1}(B')$,

f) $f^{-1}(A' \cap B') = f^{-1}(A') \cap f^{-1}(B')$,

g) $f^{-1}(A' \cup B') = f^{-1}(A') \cup f^{-1}(B')$;

如果 $Y \supset A' \supset B'$, 那么

h) $f^{-1}(A' \backslash B') = f^{-1}(A') \backslash f^{-1}(B')$,

i) $f^{-1}(C_Y A') = C_X f^{-1}(A')$;

对任何集 $A \subset X$ 及任何集 $B' \subset Y$:

j) $f^{-1}(f(A)) \supset A$,

k) $f(f^{-1}(B')) \subset B'$.

4. 试证映射 $f: X \to Y$ 是

a) 满射, 当且仅当对任何 $B' \subset Y$ 有 $f(f^{-1}(B')) = B'$.

b) 双射, 当且仅当对任何 $A \subset X$ 及任何 $B' \subset Y$ 有

$$(f^{-1}(f(A)) = A) \wedge (f(f^{-1}(B')) = B').$$

5. 检验关于映射 $f: X \to Y$ 的诸命题的等价性:

a) f 是内射;

b) 对于任意集合 $A \subset X$, 有 $f^{-1}(f(A)) = A$;

c) 对 X 的任意两个子集 A、B, 有 $f(A \cap B) = f(A) \cap f(B)$;

d) $f(A) \cap f(B) = \varnothing \Leftrightarrow A \cap B = \varnothing$;

e) 当 $X \supset A \supset B$ 时, 有 $f(A \backslash B) = f(A) \backslash f(B)$.

6. a) 如果映射 $f: X \to Y$ 与 $g: Y \to X$ 能使 $g \circ f = e_X$, 这里 e_X 是集 X 的恒等映射, 则称 g 是 f 的*左逆映射*, 而称 f 是 g 的*右逆映射*. 试证, 与逆映射的唯一性不同, 可能存在许多单边逆映射.

例如, 可考虑映射 $f: X \to Y$ 和 $g: Y \to X$, 这里 X 是单元素集, 而 Y 是二元素集, 或考虑序列的映射

$$(x_1, \cdots, x_n, \cdots) \xrightarrow{f_a} (a, x_1, \cdots, x_n, \cdots)$$
$$(y_2, \cdots, y_n, \cdots) \xleftarrow{g} (y_1, y_2, \cdots, y_n, \cdots).$$

b) 设 $f: X \to Y$ 与 $g: Y \to Z$ 都是双射. 试证, 映射 $g \circ f: X \to Z$ 是双射, 并且 $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

c) 试证, 对任何映射 $f: X \to Y, g: Y \to Z$, 及任何集合 $C \subset Z$, 等式

$$(g \circ f)^{-1}(C) = f^{-1}(g^{-1}(C))$$

成立.

d) 验证由 $(x,y)\mapsto(y,x)$ 给出的映射 $F: X\times Y\to Y\times X$ 是双射. 描述互逆映射 $f: X\to Y$ 与 $f^{-1}: Y\to X$ 之间的图像的相互联系.

7. a) 试证对任何映射 $f: X\to Y$ 来说, 由对应 $x\overset{F}{\mapsto}(x,f(x))$ 所确定的映射 $F: X\to X\times Y$ 是内射.

b) 设质点在圆周 Y 上作匀速运动; X 是时间轴, 而 $x\overset{f}{\mapsto}y$ 是瞬时 $x\in X$ 与质点位置 $y=f(x)\in Y$ 间的对应关系. 绘出函数 $f: X\to Y$ 在 $X\times Y$ 中的图像.

8. a) 对 §3 中所选的例题 1~12, 说明其中的映射是不是满射, 单射, 双射, 或者不是上述的任何一种.

b) 欧姆定律$I=V/R$ 把导线中的电流强度 I, 导线两端间的电压 V 及导线的电阻 R 联系起来. 试问, 怎样的集合间的映射 $O: X\to Y$ 对应于欧姆定律? 怎样的集合的子集对应着欧姆定律的关系?

c) 试求伽利略变换与洛伦兹变换的逆变换 G^{-1}, L^{-1}.

9. a) 设有集合 $S\subset X$ 及映射 $f: X\to X$. 如果 $f(S)\subset S$, 则称 S 关于 f 是*稳定的*. 写出关于平面沿其上一向量的平移的稳定集.

b) 设集 $I\subset X$, 映射 $f: X\to X$. 如果 $f(I)=I$, 则称 I 关于 f 是*不变的*. 写出关于平面绕一个定点旋转的不变集.

c) 点 p 叫做映射$f: X\to X$ 的*不动点*, 如果 $f(p)=p$. 试证平面上的平移、旋转与位似的任意复合, 当位似系数小于 1 时, 有不动点.

d) 把伽利略变换及洛伦兹变换看做映平面到自身上的映射, 它将坐标为 (x,t) 的点变到坐标为 (x',t') 的点. 试求这些变换的不变集.

10. 考察液体的稳态流动 (即每个点的流速不随时间变化). 在时间 t 内, 位于点 x 处的质点变到了空间某一新位置 $f_t(x)$. 这样产生的由流体占据的空间的点的映射 $x\mapsto f_t(x)$ 与时间有关, 称之为*时间 t 内的变换*. 试证

$$f_{t2}\circ f_{t1}=f_{t1}\circ f_{t2}=f_{t1+t2}\text{和}f_t\circ f_{-t}=f_0=e_X.$$

§4. 某些补充

1. 集的势 (基数) 设 X,Y 为二集合. 如果存在 X 到 Y 上的双射, 即对每个 $x\in X$, 有 $y\in Y$ 与之对应, X 中不同的元, 在 Y 中所对应的元也不同, 并且每个 $y\in Y$, 必是 X 中某元的对应元, 则称 X 与 Y等势.

通俗地说, 就是每个元 $x\in X$ 坐在自己的位子 $y\in Y$ 上, X 的所有元都坐着, 而且没有闲置的位子 $y\in Y$.

显然, 所引入的关系 $X\mathcal{R}Y$ 是等价关系, 按照约定, 我们在这种情况把 $X\mathcal{R}Y$ 写成 $X\sim Y$.

等势关系把所有集合所成的族分成了彼此等价的集合的类. 同一个等价类中的集, 有相同数量的元素 (等势), 不同类中的集合所含元素数量不同.

集 X 所在的类叫集 X 的势, 或叫集 X 的基数, 并记作 $\operatorname{card}X$. 如果 $X \sim Y$, 就写成 $\operatorname{card}X = \operatorname{card}Y$.

这种结构的意义在于它使我们能够比较集合的元素的数量, 而不使用数数这个中间步骤, 即不必用与自然数列 $\mathbb{N} = \{1, 2, 3, \cdots\}$ 对照的方法确定数量; 很快我们就会看到, 这后一种方法, 有时根本不能用.

如果集合 X 与集合 Y 的某个子集等势, 我们就说 X 的基数不大于 Y的基数, 并记作 $\operatorname{card}X \leqslant \operatorname{card}Y$.

因此,

$$(\operatorname{card}X \leqslant \operatorname{card}Y) := (\exists Z \subset Y | \operatorname{card}X = \operatorname{card}Z).$$

如果 $X \subset Y$, 则显然 $\operatorname{card}X \leqslant \operatorname{card}Y$. 然而关系 $X \subset Y$ 却不妨碍有不等式 $\operatorname{card}Y \leqslant \operatorname{card}X$, 即便 X 是 Y 的真子集, 也是如此.

例如, 对应 $x \mapsto \dfrac{x}{1-|x|}$ 是数轴 $\mathbb{R}$ 的开区间 $-1 < x < 1$ 到整个数轴本身的双射.

一集合能与其自己的部分等势, 是这个集合为无穷集的特征标志, 戴德金[①] 甚至曾建议把它作为无穷集的定义. 这样, 一个集合 (按照戴德金的说法), 只有当它不与自己的任何真子集等势时叫做*有穷集*, 不然就叫*无穷集*.

就像不等关系把数轴上的实数有序化一样, 上边引进的不等关系也把集合的势或基数有序化. 这就是说, 可以证明所建立的不等关系有下列性质:

$1°$ $(\operatorname{card}X \leqslant \operatorname{card}Y) \wedge (\operatorname{card}Y \leqslant \operatorname{card}Z)$
$\Rightarrow (\operatorname{card}X \leqslant \operatorname{card}Z)$(显然);

$2°$ $(\operatorname{card}X \leqslant \operatorname{card}Y) \wedge (\operatorname{card}Y \leqslant \operatorname{card}X)$
$\Rightarrow (\operatorname{card}X = \operatorname{card}Y)$(施略德–伯恩斯坦[②] 定理);

$3°$ $\forall X \forall Y (\operatorname{card}X \leqslant \operatorname{card}Y) \vee (\operatorname{card}Y \leqslant \operatorname{card}X)$(康托尔定理).

因此, 基数类是有线性序的.

如果 $\operatorname{card}X \leqslant \operatorname{card}Y$, 同时 $\operatorname{card}X \neq \operatorname{card}Y$, 就说集 X 的势小于集 Y 的势, 并记作 $\operatorname{card}X < \operatorname{card}Y$. 于是,

$$(\operatorname{card}X < \operatorname{card}Y) := (\operatorname{card}X \leqslant \operatorname{card}Y) \wedge (\operatorname{card}X \neq \operatorname{card}Y).$$

像前面一样, 用 $\varnothing$ 记空集, 而用 $\mathcal{P}(X)$ 记 X 的一切子集构成的集. 康托尔发现有以下的

---

①戴德金(Dedekind)(1831—1916)—— 德国数学家, 他积极参与了发展实数理论的工作, 首先提出自然数集的公理化系统, 通常它用意大利数学家佩亚诺 (Peano, 1858—1932) 的名字命名为佩亚诺公理系统. 佩亚诺在戴德金提出这个公理系统稍后系统地叙述了这个公理系统.

②F.伯恩斯坦(F. Bernstein)(1878—1956)—— 德国数学家, 康托尔的学生; 施略德(E. Schröder)(1841—1902)—— 德国数学家.

定理 (康托尔) $\mathrm{card}X < \mathrm{card}\mathcal{P}(X)$.

◀ 对于空集 $\varnothing$ 来说, 上述结论显然成立, 所以以下可认为 $X \neq \varnothing$.

因为 $\mathcal{P}(X)$ 含有 X 的一切单元素子集, 所以, $\mathrm{card}X \leqslant \mathrm{card}\mathcal{P}(X)$.

现在只需证明, 当 $X \neq \varnothing$ 时, $\mathrm{card}X \neq \mathrm{card}\mathcal{P}(X)$.

如果不如此, 假定 $f: X \to \mathcal{P}(X)$ 是双射. 考察集合 $A = \{x \in X | x \notin f(x)\}$, 它由那样一些元素 $x \in X$, x 不含于它所对应的集 $f(x) \in \mathcal{P}(X)$, 组成的. 因为 $A \in \mathcal{P}(X)$, 所以必能找到一个元素 $a \in X$, 使得 $f(a) = A$. 这个元素 $a \in X$ 既不能有 $a \in A$(据 A 的定义), 也不能有 $a \notin A$(还是根据 A 的定义). 这与排中律矛盾 ▶

特别地, 这个定理说明, 如果无穷集存在, 那么,"无穷" 必是各种各样的.

2. 公理化集合论 本段目的是为有兴趣的读者提供集合公理系统概念, 这些公理描述了称之谓集合的数学对象的性质, 并证明这些公理的一些最简单的推论.

1° **容积公理** *集合 A 与集合 B 相等, 当且仅当它们有共同的元素.*

这就是说, 我们并不去管对象 "集合" 的一切其他性质, 只注意它的元素有哪些就行了. 在实际应用时就是, 如果要验证 $A = B$, 只需要验证 $\forall x((x \in A) \Leftrightarrow (x \in B))$.

2° **分出公理** *对任何集合 A 及性质 P, 有这样的集 B, 它所含的元素, 是且仅是 A 中的那些具有性质 P 的元素.*

简短地说, 如果 A 是一个集合, 那么, 可以断定 $B = \{x \in A | P(x)\}$ 也是一个集合.

当从一个集合分出另一个具有这种或那种性质的子集, 以便作出数学结构时, 经常要用到这个公理.

例如, 由分出公理推知, 任何集合 X 有空子集$\varnothing_X = \{x \in X | x \neq x\}$, 再注意到容积公理, 就知道, 对于任意集合 X 与 Y, 有 $\varnothing_X = \varnothing_Y$, 即空集是唯一的. 用符号 $\varnothing$ 表示空集.

由分出公理还推知, 如果 A, B 都是集合, 那么 $A \backslash B = \{x \in A | x \notin B\}$ 也是集合. 特别当 M 是一个集合而 A 是它的子集时, $C_M A$ 也是一个集合.

3° **并公理** *对于由集合构成的任何集合 M, 存在集合 $\cup M$, 称为集合 M 的并, 它的元素恰好是 M 中所含元素的元素.*

如果把 "集合的集合" 换成说 "集合的族", 那么, 并公理就获得了更加习惯的说法: 存在着由族中诸集合的元素组成的集合. 这样, 集合的并仍是一个集合, 并且 $x \in \cup M \Leftrightarrow \exists X((X \in M) \wedge (x \in X))$.

并公理再加上分出公理, 就能把 (集族) M 的交定义成是一个集

$$\cap M := \{x \in \cup M | \forall X((X \in M) \Rightarrow (x \in X))\}.$$

4° **对公理** *对于任意集合 X 与 Y, 存在一个集合 Z, 使 X 与 Y 是它仅有的元素.*

集合 Z 用 $\{X, Y\}$ 表示, 并称之为集合 X 与 Y 的无序对. 如果 $X = Y$, 那么集合 Z 只有一个元素.

我们已经知道, 集合的序对(X, Y) 与无序对的差异在于对其中一个集合加了某种标志. 例如,

$$(X, Y) := \{\{X, X\}, \{X, Y\}\}.$$

因此, 能用无序对引进序对; 如果利用分出公理及下面的重要公理, 就能由序对引进集的直积.

5° 子集之集的公理 *对于任意集合 X, 存在集合 $\mathcal{P}(X)$, 它的元素恰好就是 X 的一切子集.*

简短地说, 存在着以已知集合的一切子集为元素的集合.

现在能验证序对 (x, y), 其中 $x \in X, y \in Y$, 确实构成集合

$$X \times Y := \{p \in \mathcal{P}(\mathcal{P}(X) \cup \mathcal{P}(Y)) | p = (x, y) \wedge (x \in X) \wedge (y \in Y)\}.$$

公理 1° ~ 5° 限制了形成新集的可能, 譬如, 在集合 $\mathcal{P}(X)$ 中有不属于 X 的元素 (这因为根据康托尔定理: $\mathrm{card}X < \mathrm{card}\mathcal{P}(X)$), 所以不存在所有集之"集". 须知, 这正是罗素悖论立足之"集"!

为了叙述下面的公理, 我们引进一个集合 X 的*后续集* X^+ 概念. 定义 $X^+ = X \cup \{X\}$, 简言之, 就是在 X 上添加一个单元素集 $\{X\}$.

此外, 我们称集合为*归纳集*, 假如它包含空集以及其中任何一集的后续集.

6° 无穷公理 *归纳集存在*

根据无穷性公理, 连同公理 1° ~ 4°,(按冯 · 诺依曼 (Neumann) ① 的建议) 可以建立自然数集 $\mathbb{N}_0$ 的标准模型, 方法是把 $\mathbb{N}_0$ 定义为一切归纳集的交, 即最小归纳集. $\mathbb{N}_0$ 的元素是集合

$$\varnothing, \quad \varnothing^+ = \varnothing \cup \{\varnothing\} = \{\varnothing\}, \quad \{\varnothing\}^+ = \{\varnothing\} \cup \{\{\varnothing\}\}, \quad \cdots,$$

它们就是我们用符号 0,1,2, ··· 表示并称之为自然数的那些东西的模型.

7° 置换公理 *设 $\mathcal{F}(x, y)$ 是这样的一个命题 (更确切地说, 是公式), 使得对于集 X 中的任何 x_0, 存在唯一的对象 y_0, 使得 $\mathcal{F}(x_0, y_0)$ 成立. 这时, 存在 $x \in X$, 使得 $\mathcal{F}(x, y)$ 成立的那些 y 能组成集合.*

建立分析学时, 我们并不用这一公理.

公理 1° ~ 7° 组成了集合论公理系统, 即著名的策梅洛-弗兰克尔 (Zermelo- Fraenkel)② 公理系统.

对于它们, 通常还要增加一个与公理 1° ~ 7° 独立, 且在分析中经常用到的

8° 选择公理 *对于任意不空集的族, 存在一集合 C, 使对所给族中的每个集合 X, 集合 $X \cap C$ 恰好只含一个元素.*

换句话说, 恰好可从族的每个集中选出一个代表, 由它们组成集合 C.

选择公理, 即数学中著名的策梅洛公理, 在当时引起了专家们的激烈的争论.

3. 关于数学命题的逻辑结构及其用集合论语言的写法的注记 在集合论的语言中, 有两个基本的或者说是原子型的数学命题: 一个是 $x \in A$, 它是说客体 x 是集合 A 的元素; 另一个是 $A = B$, 它是说集合 A 与集合 B 相同. (可是, 注意到容积公理, 就知第二个断语是第一型断语的组合 $((x \in A) \Leftrightarrow (x \in B))$).

复杂的命题或复杂的逻辑式, 由原子型的命题借助逻辑运算并利用括号 ( ) 构成, 这些运算是: 连接词 $\neg, \wedge, \vee, \Rightarrow$, 和量词 $\forall, \exists$. 同时, 无论多么复杂的命题的表示及其写法, 都可归结为完成下面的一些初等逻辑运算:

---

①冯 · 诺依曼 (1903—1957)—— 美国数学家. 他从事泛函分析, 量子力学的数学基础, 拓扑群, 博弈论和数理逻辑研究工作, 领导建造了第一架电子计算机.

②策梅洛(Zermelo, E. F. F.) (1871 -1953)是德国数学家; Fraenkel, A (1891 1965)是德国, 后是以色列数学家.

a) 在某些命题之前放上否定号并把所得结果用括号括起来作成新命题.

b) 在两个命题间放上必要的连接词 $\wedge, \vee, \Rightarrow$ 等, 再把结果括起来作成新命题.

c) 构成命题 "对任何 x 都有性质 P"(这句话写成 $\forall x P(x)$) 或构成命题 "存在具有性质 P 的 x".(写成 $\exists x P(x)$).

例如

$$\exists x((P(x)) \wedge (\forall y((P(y)) \Rightarrow (y = x)))),$$

就是说, 存在那样的客体 x, 它具有性质 P, 并且, 如果 y 是具有性质 P 的任何客体, 则 $y = x$. 简言之, 这就是说存在唯一客体 x 具有性质 P. 通常把这一命题简短地表示作 $\exists! x P(x)$, 我们以后也利用这一缩记法.

为了缩短命题的写法, 正像已经看到的那样, 要在不丧失写法一意性的前提下, 尽量减少括号的数量; 为此, 除了早先指出的运算 $\neg, \wedge, \vee, \Rightarrow$ 的优先原则之外, 还认为公式中最占先的是 $\in, =$, 其次是 $\exists, \forall$; 最后才是 $\neg, \wedge, \vee, \Rightarrow$

作了这些约定之后, 就可以写出:

$$\exists! x P(x) := \exists x(P(x) \wedge \forall y(P(y) \Rightarrow y = x)).$$

关于以下广泛使用的缩记法, 我们还约定:

$$\begin{aligned}
(\forall x \in X)P &:= \forall x(x \in X \Rightarrow P(x)),\\
(\exists x \in X)P &:= \exists x(x \in X \wedge P(x)),\\
(\forall x > a)P &:= \forall x(x \in \mathbb{R} \wedge x > a \Rightarrow P(x)),\\
(\exists x > a)P &:= \exists x(x \in \mathbb{R} \wedge x > a \wedge P(x)).
\end{aligned}$$

在这里, 如通常一样, $\mathbb{R}$ 表示实数集.

利用这些缩记法和构造复杂命题的规则 a),b),c), 可以用

$$\begin{aligned}
(\lim_{x \to a} f(x) = A) := \forall \varepsilon > 0 \exists \delta > 0 \forall x \in \mathbb{R}(0 < |x - a| < \delta\\
\Rightarrow |f(x) - A| < \varepsilon)
\end{aligned}$$

来解释数 A 是函数 $f : \mathbb{R} \to \mathbb{R}$ 在点 $a \in \mathbb{R}$的极限.

本节所说的一切, 对我们最重要的, 可能是以下关于构造含有量词的命题的否命题的法则.

命题 "有 x 使 $P(x)$ 成立" 的否定, 就是 "无论 x 如何, $P(x)$ 不成立"; 而命题 "对于任何 $x, P(x)$成立" 的否定, 则是 "存在 x 使 $P(x)$ 不成立".

于是

$$\neg \exists x P(x) \Leftrightarrow \forall x \neg P(x),$$

$$\neg \forall x P(x) \Leftrightarrow \exists x \neg P(x).$$

我们还应注意 (参看 §1 的习题)

$$\neg(P \wedge Q) \Leftrightarrow \neg P \vee \neg Q,$$
$$\neg(P \vee Q) \Leftrightarrow \neg P \wedge \neg Q,$$
$$\neg(P \Rightarrow Q) \Leftrightarrow P \wedge \neg Q.$$

基于所说的这些, 即可推知, 比如

$$\neg((\forall x > a)P) \Leftrightarrow (\exists x > a)\neg P.$$

当然, 如果把上式右端写成 $(\exists x \leqslant a)\neg P$ 那就错了.

事实上,

$$\begin{aligned}\neg((\forall x > a)P) : &= \neg(\forall x(x \in \mathbb{R} \wedge x > a \Rightarrow P(x)))\\ &\Leftrightarrow \exists x \neg(x \in \mathbb{R} \wedge x > a \Rightarrow P(x))\\ &\Leftrightarrow \exists x((x \in \mathbb{R} \wedge x > a) \wedge \neg P(x)) =: (\exists x > a)\neg P.\end{aligned}$$

如果看一看上面所指出的任何命题的结构, 那么, 借助于已经得到的这些最简单命题的否定, 就能够把任何一个具体命题的否定构造出来.

例如

$$\neg(\lim_{x\to a} f(x) = A) \Leftrightarrow \exists \varepsilon > 0 \forall \delta > 0 \exists x \in \mathbb{R}(0 < |x-a| < \delta \wedge |f(x) - A| \geqslant \varepsilon).$$

特别地, 正确构造否命题在实践上的重要性, 是与反证法密切联系的, 那时要从断言 $\neg P$ 不成立推出断言 P 成立.

练 习

1. a) 利用施略德–伯恩斯坦定理及直接表出双射关系两种方法, 证明直线上的闭区间 $\{x \in \mathbb{R}|0 \leqslant x \leqslant 1\}$ 与开区间 $\{x \in \mathbb{R}|0 < x < 1\}$ 等势.

b) 分析施略德–伯恩斯坦定理

$$(\mathrm{card}X \leqslant \mathrm{card}Y) \wedge (\mathrm{card}Y \leqslant \mathrm{card}X) \Rightarrow (\mathrm{card}X = \mathrm{card}Y)$$

的下述证明.

◀ 只需证明: 如果集合 X, Y, Z 满足 $X \supset Y \supset Z$ 且 $\mathrm{card}X = \mathrm{card}Z$, 则 $\mathrm{card}X = \mathrm{card}Y$. 设 $f : X \to Z$ 是双射. 那么, 可用下面的方式给出双射 $g : X \to Y$

$$g(x) = \begin{cases} f(x), \text{当有某个} n \text{使} x \in f^n(X)\backslash f^n(Y), \\ x, \quad \text{其他情形}. \end{cases}$$

这里 $f^n = f \circ \cdots \circ f$ 是映射 f 的 n 次迭代, 而 $\mathbb{N}$ 是自然数集. ▶

2. a) 从对的定义出发, 验证在第二段中所下的二集合 X, Y 的直积 $X \times Y$ 的定义是正确的, 即集 $\mathcal{P}(\mathcal{P}(X) \cup \mathcal{P}(Y))$ 包含着一切序对, 其中 $x \in X, y \in Y$.

b) 试证, 由一个固定的集合 X, 到另一个固定的集合 Y 内的一切可能的映射 $f: X \to Y$ 本身也组成集合 $M(X, Y)$.

c) 验证, 如果 $\mathcal{R}$ 是序对 (即是关系) 的集合, 那么, 属于集 $\mathcal{R}$ 的序对的第一个元素 (第二个元素) 自身构成集合.

3. a) 利用容积公理, 对公理, 分出公理, 并公理与无穷公理, 验证对自然数集 $\mathbb{N}_0$ 的元素 (据冯 · 诺依曼), 以下的断语是正确的:

1° $x = y \Rightarrow x^+ = y^+$;

2° $(\forall x \in \mathbb{N}_0)(x^+ \neq \varnothing)$;

3° $x^+ = y^+ \Rightarrow x = y$;

4° $(\forall x \in \mathbb{N}_0)(x \neq \varnothing \Rightarrow (\exists y \in \mathbb{N}_0)(x = y^+))$.

b) 利用 $\mathbb{N}_0$ 是归纳集, 证明对 $\mathbb{N}_0$ 的任何元素 x, y (而它们本身是集合) 来说, 以下诸关系成立:

1° $\text{card}\, x \leqslant \text{card}\, x^+$;

2° $\text{card}\, \varnothing < \text{card}\, x^+$;

3° $\text{card}\, x < \text{card}\, y \Leftrightarrow \text{card}\, x^+ < \text{card}\, y^+$;

4° $\text{card}\, x < \text{card}\, x^+$;

5° $\text{card}\, x < \text{card}\, y \Rightarrow \text{card}\, x^+ \leqslant \text{card}\, y$;

6° $x = y \Leftrightarrow \text{card}\, x = \text{card}\, y$;

7° $(x \subset y) \vee (x \supset y)$.

c) 试证, 在集 $\mathbb{N}_0$ 的任何子集 X 中存在这样的 (最小) 元素 x_m, 使得 $(\forall x \in X)(\text{card}\, x_m \leqslant \text{card}\, x)$. (如果做这个题有困难, 可以在学完第二章后再做).

4. 我们将只涉及集合. 既然由各种元素组成的集合, 本身又可能是其他集合的元素, 在逻辑上, 所有集合通常都用小写字母来记, 这种记法在这个题目中是很方便的.

a) 验证写法

$$\forall x \exists y \forall z (z \in y \Leftrightarrow \exists w (z \in w \wedge w \in x))$$

表示并公理, 据此公理, 作为集合 x 的并的集合 y 是存在的.

b) 请指出以下所写诸式各表示集合论的什么公理

$$\forall x \forall y \forall z ((z \in x \Leftrightarrow z \in y) \Leftrightarrow x = y),$$

$$\forall x \forall y \exists z \forall v (v \in z \Leftrightarrow (v = x \vee v = y)),$$

$$\forall x \exists y \forall z (z \in y \Leftrightarrow \forall u (u \in z \Rightarrow u \in x)),$$

$$\begin{aligned} \exists x (\forall y (\neg \exists z (z \in y) \Rightarrow y \in x) \wedge \forall w (w \in x \\ \Rightarrow \forall u (\forall v (v \in u \Leftrightarrow (v = w \vee v \in w)) \\ \Rightarrow u \in x))). \end{aligned}$$

c) 验证公式

$$\begin{aligned}&\forall z(z \in f \Rightarrow (\exists x_1 \exists y_1 (x_1 \in x \wedge y_1 \in y \wedge z = (x_1, y_1)))) \wedge \\ &\wedge \forall x_1 (x_1 \in x \Rightarrow \exists y_1 \exists z (y_1 \in y \wedge z = (x_1, y_1) \wedge z \in f)) \wedge \\ &\wedge \forall x_1 \forall y_1 \forall y_2 (\exists z_1 \exists z_2 (z_1 \in f \wedge z_2 \in f \wedge z_1 = (x_1, y_1) \wedge \\ &\wedge z_2 = (x_1, y_2)) \Rightarrow y_1 = y_2).\end{aligned}$$

逐步对集合 f 加上三条约束:f 是 $x \times y$ 的子集; f 在 x 上的射影与 x 重合; x 的每个元素 x_1, 恰好对应着 y 中的一个元素 y_1, 使 $(x_1, y_1) \in f$.

因此, 定义了一个映射 $f : x \to y$.

这个例子再一次说明: 形式化书写的命题, 绝不总是比口语更简短更明显. 鉴于这种情况, 我们在以后, 只是在能使叙述更加紧凑或更加明显时, 才使用逻辑符号.

5. 设 $f : X \to Y$ 是一个映射, 试写出下面每个命题的逻辑否定式:

a) f 是满射;

b) f 是内射;

c) f 是双射.

6. 设 X 与 Y 都是集合, $f \subset X \times Y$, 试写出一个不是函数的集 f 来.

第二章　实数

一般来说, 数学理论是在那些能把原来给出的一套数据加工成中间计算目标或最终计算目标所需要的另一套数据的地方寻找自己的出路的. 由于这个缘故, 数值函数在数学及其应用中占有特殊的地位, 它们 (确切地说, 是所谓可微数值函数) 构成了古典分析的主要研究对象. 但是, 要对你们在中学里已经接触到并且不久将更加深信的这些函数的性质作一个从现代数学观点看稍微完整的描述, 若不先精确地定义使这些函数起作用的实数集, 那是不可能的.

数学中的数, 就像物理中的时间一样, 人人都知道, 唯独专家们不这样理解它. 这是数学基本抽象之一. 看来, 它仍面临本质的演变. 关于它的故事可能要专门写一本大部头的书. 在这里, 我们只把基本上是读者在中学里学过的有关实数的知识加以总结, 并以公理的形式抽出那些基本的和独立的数的性质. 同时我们的目的在于给出一个确切的而且在以后用起来方便的实数定义, 并特别注意于它的完备性或连续性, 因为它们是分析中基本的非算术运算, 即极限过程的胚芽.

§1. 实数集的公理系统及它的某些一般性质

1. 实数集的定义

定义 1 满足以下四组条件的集 $\mathbb{R}$ 叫实数集, 它的元素叫实数, 这些条件一起构成实数集的公理系统:

(I) 加法公理 确定了一个映射 (加法运算)

$$+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R},$$

使得对于 $\mathbb{R}$ 中任意二元 x, y 之序对 (x, y), 有某元 $x+y \in \mathbb{R}$ 与之对应, 称 $x+y$ 为 x、y 之和, 同时此映射满足以下条件:

1_+. 有中性元0 存在 (叫做加法零元), 使对任何的 $x \in \mathbb{R}$,

$$x+0=0+x=x.$$

2_+. 每个元 $x \in \mathbb{R}$ 有元 $-x \in \mathbb{R}$, 叫做 x 的负元, 使得

$$x+(-x)=(-x)+x=0.$$

3_+. 运算 $+$ 是结合的, 即 $\mathbb{R}$ 中任何 x, y, z 满足条件

$$x+(y+z)=(x+y)+z.$$

4_+. 运算 $+$ 是交换的, 即 $\mathbb{R}$ 中的任何 x, y 满足

$$x+y=y+x.$$

如果在任何集 G 上确定了满足公理 $1_+, 2_+, 3_+$ 的一个运算, 就说在 G 上给定了一个群结构, 或说 G是群. 如果称此运算为加法, 那么就称这个群为*加群*. 如果这运算又是交换的, 即满足条件 4_+, 那么这个群叫做*交换群* 或阿贝尔群[①].

于是, 公理 1_+—4_+ 说明, $\mathbb{R}$ 是阿贝尔群.

(II) 乘法公理 确定了一个映射 (乘法运算)

$$\bullet : \mathbb{R} \times \mathbb{R} \to \mathbb{R},$$

使得对于 $\mathbb{R}$ 中任意二元 x、y 之序对 (x, y), 有某元 $x \cdot y \in \mathbb{R}$ 与之对应, 称 $x \cdot y$ 为 x 与 y 之积, 并且满足以下条件:

$1_\bullet$. 存在中性元$1 \in \mathbb{R}\backslash 0$(乘法的中性元叫单位元) 使得 $\forall x \in \mathbb{R}$,

$$x \cdot 1 = 1 \cdot x = x.$$

$2_\bullet$. 每个元 $x \in \mathbb{R}\backslash 0$ 有元 $x^{-1} \in \mathbb{R}\backslash 0$, 叫做 x 的逆元, 满足

$$x \cdot x^{-1} = x^{-1} \cdot x = 1.$$

$3_\bullet$. 运算 $\bullet$ 是结合的, 即任何 $x, y, z \in \mathbb{R}$ 满足

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z.$$

$4_\bullet$. 运算 $\bullet$ 是交换的, 即任何 $x, y \in \mathbb{R}$ 满足

$$x \cdot y = y \cdot x.$$

可以验证, 集 $\mathbb{R}/0$ 关于乘法运算是一个 (乘法) 群.

---

①阿贝尔(N. H. Abel)(1802—1829)—— 著名挪威数学家, 他证明了高于四次的代数方程不能用根式解出.

(I,II) 加法与乘法的联系 乘法对加法有分配性,即 $\forall x, y, z \in \mathbb{R}$,

$$(x+y)\cdot z = x\cdot z + y\cdot z.$$

注意, 由于乘法有可交换性, 最后这个等式, 在改变等号两端乘法顺序后, 仍然成立.

如果集 G 上定义了两种运算, 它们满足上面所有的公理, 就称 G 是一个代数域或简称域.

(III) 序公理 $\mathbb{R}$的元素间有关系 $\leqslant$, 即对 $\mathbb{R}$ 的元素 x 与 y, 或满足 $x \leqslant y$, 或不满足. 同时应满足以下条件:

$0_{\leqslant}$. $\forall x \in \mathbb{R}(x \leqslant x)$.

$1_{\leqslant}$. $(x \leqslant y) \land (y \leqslant x) \Rightarrow (x = y)$.

$2_{\leqslant}$. $(x \leqslant y) \land (y \leqslant z) \Rightarrow (x \leqslant z)$.

$3_{\leqslant}$. $\forall x \in \mathbb{R} \forall y \in \mathbb{R}(x \leqslant y) \lor (y \leqslant x)$.

$\mathbb{R}$ 中的关系 $\leqslant$ 叫做*不等关系*.

大家都知道, 如果某个集合的某些元素之间有满足公理 $0_{\leqslant}, 1_{\leqslant}, 2_{\leqslant}$ 的关系, 就称这集合为*偏序集*; 如果除此之外又满足 $3_{\leqslant}$, 即集合中的任二元素均能比较大小, 就称此集合为*线性序集*.

因此, 实数集对于它的元素间的不等关系来说, 是线性序集.

(I,III) $\mathbb{R}$ 中的加法与序关系的联系 如果 x, y, z 是 $\mathbb{R}$ 的元素, 那么,

$$(x \leqslant y) \Rightarrow (x+z \leqslant y+z).$$

(II,III) $\mathbb{R}$ 中的乘法与序关系的联系 如果 x, y 是 $\mathbb{R}$ 的元素, 那么,

$$(0 \leqslant x) \land (0 \leqslant y) \Rightarrow (0 \leqslant x\cdot y).$$

(IV) 完备 (连续)公理 如果 X 与 Y 是 $\mathbb{R}$ 的非空子集, 且具有性质: 对于任何 $x \in X, y \in Y$, 有 $x \leqslant y$, 那么, 存在 $c \in \mathbb{R}$, 使对任何 $x \in X, y \in Y$ 有 $x \leqslant c \leqslant y$.

这些就是公理一览表, 满足这些公理的任何集合 $\mathbb{R}$, 都可以被认为是实数集的具体实现, 或通常所说的*实数模型*.

这个定义形式上并不假定事先知道关于数的任何知识, 应当做的是 "开动数学思维的机器", 从这个定义出发, 把实数的其他性质作为定理, 形式地推导出来. 对于这种公理形式主义, 我们想作些非形式主义的说明.

设想我们没有经过数苹果, 玩积木或其他具体导致抽象自然数加法的数数阶段; 我们没有度量过线段, 从而不曾接触过有理数; 我们不知道古代关于正方形对角线与其边无公度的伟大发现, 从而也不知道 (单位边长的正方形的) 对角线的长不能是有理数, 即引进无理数的必要性; 我们没有产生于测量过程中的 "大于" ("小于") 概

念; 我们不能, 比如说, 用数轴的形象来解释顺序关系. 假如所有这些, 我们事先真不知道的话, 那么所列出的这一套公理, 就不仅不能作为思想发展的一个总结为我们所接受, 而且至少显得古怪, 从而完全成为一种臆造出来的东西.

对于任何抽象公理系统, 至少有两个问题, 马上会想到.

第一, 这组公理是否相容, 即是否存在着满足列出的所有条件的集合. 这是*公理系统的无矛盾性*问题.

第二, 这组公理所确定的数学对象是否是唯一的, 即如逻辑学所说, 这组公理是不是*范畴的*. 在这里, 对唯一性必须作如下理解. 比如 A 和 B 两人彼此独立地建立了满足公理的数系模型 $\mathbb{R}_A$ 和 $\mathbb{R}_B$, 那么集合 $\mathbb{R}_A$ 与 $\mathbb{R}_B$ 间可以建立双射 $f:\mathbb{R}_A\to\mathbb{R}_B$, 它保持算术运算与序关系, 即

$$f(x+y)=f(x)+f(y),$$
$$f(x\cdot y)=f(x)\cdot f(y),$$
$$x\leqslant y\Leftrightarrow f(x)\leqslant f(y).$$

这时, 从数学的观点来看, $\mathbb{R}_A$ 与 $\mathbb{R}_B$ 只是实数的不同的 (完全平等的) 实现 (模型) (例如 $\mathbb{R}_A$ 是无穷十进小数, 而 $\mathbb{R}_B$ 是数轴上的点). 这些实现叫做*同构实现*, 而映射 f 则叫做*同构*. 这样, 数学研究的结果, 就不只适用于个别的实现, 而且对该公理系统的同构模型类中每个模型都是适用的.

我们这里不去讨论上面提出的问题, 而只限于叙述它的答案.

关于无矛盾性公理问题的答案, 总带有一些假定性的特点. 对于数来说, 大致是这样做的: 从被我们所接受的集合论公理系统出发 (参看第一章 §4 第 2 段) 可建立自然数集, 然后是有理数集, 最后是满足以上所有性质的整个实数集合.

至于实数公理系统的范畴性问题, 答案也是肯定的. 有兴趣的读者能自己得到这个结果, 为此只要做下节末的练习 23,24.

2. 实数的某些一般的代数性质 我们用例子说明, 那些众所周知的数的性质是怎样从上面引进的公理得出来的.

a. 加法公理的推论

1° *实数集中有唯一的零元.*

◀ 如果 0_1 与 0_2 都是 $\mathbb{R}$ 中的零, 由零的定义得到

$$0_1=0_1+0_2=0_2+0_1=0_2. ▶$$

2° *实数集中的每个元素有唯一的负元.*

◀ 如果 x_1 与 x_2 都是 $x\in\mathbb{R}$ 的负元, 那么,

$$x_1=x_1+0=x_1+(x+x_2)=(x_1+x)+x_2$$
$$=x_2+(x+x_1)=x_2+0=x_2. ▶$$

在这里, 我们依次用了零的定义, 负元的定义, 加法的结合性, 再一次用了负元的定义, 最后又用了零的定义.

3° 方程 $a+x=b$ 在 $\mathbb{R}$ 中有唯一解 $x=b+(-a)$.

◀ 这由每个元 $a\in\mathbb{R}$ 有唯一的负元得到:

$$\begin{aligned}(a+x=b) &\Leftrightarrow ((x+a)+(-a)=b+(-a))\\ &\Leftrightarrow (x+(a+(-a))=b+(-a)) \Leftrightarrow (x+0=b+(-a))\\ &\Leftrightarrow x=b+(-a). \blacktriangleright\end{aligned}$$

式子 $b+(-a)$ 也可写成 $b-a$ 的形式. 通常我们用的就是这种简短的习惯写法.

b. 乘法公理的推论

1° 实数集中有唯一的单位元.

2° 对于每个数 $x\neq 0$, 有唯一的逆元 x^{-1}.

3° 方程 $a\cdot x=b$, 当 $a\in\mathbb{R}\backslash 0$ 时, 有唯一的解 $x=b\cdot a^{-1}$.

这些断语的证明显然是对加法相应断语证明的重复 (只要把符号和运算名称换一下就行了), 所以略去.

c. 加法与乘法联系的公理的推论 引进了联系加法与乘法的附加公理 (I,II) 之后, 就得到进一步的推论.

1° 对于任何 $x\in\mathbb{R}$

$$x\cdot 0=0\cdot x=0.$$

◀ $(x\cdot 0=x\cdot(0+0)=x\cdot 0+x\cdot 0)\Rightarrow(x\cdot 0=x\cdot 0+(-(x\cdot 0))=0)$ ▶

顺便得到, 如果 $x\in\mathbb{R}\backslash 0$, 则 $x^{-1}\in\mathbb{R}\backslash 0$.

2° $(x\cdot y=0)\Rightarrow(x=0)\vee(y=0)$.

◀ 假如 $y\neq 0$, 那么, 由关于 x 的方程 $x\cdot y=0$ 的解的唯一性得 $x=0\cdot y^{-1}=0$. ▶

3° 对于任何 $x\in\mathbb{R}$

$$-x=(-1)\cdot x.$$

◀ $x+(-1)\cdot x=(1+(-1))x=0\cdot x=x\cdot 0=0$,

再由负元的唯一性就推出所要证的论断. ▶

4° 对于任何 $x\in\mathbb{R}$

$$(-1)(-x)=x.$$

◀ 由 3° 及 $-x$ 的负元 x 的唯一性即得. ▶

5° 对于任何 $x\in\mathbb{R}$

$$(-x)(-x)=x\cdot x.$$

◀ $(-x)(-x)=((-1)\cdot x)(-x)=(x\cdot(-1))(-x)=x((-1)(-x))=x\cdot x$.

这里依次用了前面的两个断语, 还用了乘法的交换律与结合律. ▶

d. 序公理的推论 首先注意, 关系 $x \leqslant y$(读作 "x 小于或等于 y") 也可以写成 $y \geqslant x$("y 大于或等于 x"); 当 $x \neq y$ 时, 关系 $x \leqslant y$ 写成 $x < y$ (读作 "x 小于 y") 或 $y > x$ ("y 大于 x"), 并称之为*严格不等式*.

1° *对于任何* $x, y \in \mathbb{R}$, *在*

$$x < y, x = y, x > y$$

中, 恰有一种关系成立.

◀ 这从刚引入的严格不等式的定义, 以及公理 $1_{\leqslant}$ 与 $3_{\leqslant}$ 推知. ▶

2° 对于任何 $x, y, z \in \mathbb{R}$

$$(x < y) \wedge (y \leqslant z) \Rightarrow (x < z),$$
$$(x \leqslant y) \wedge (y < z) \Rightarrow (x < z).$$

◀ 比如我们来证第二个断语.

据不等式关系之传递性即公理 $2_{\leqslant}$ 推知

$$(x \leqslant y) \wedge (y < z) \Leftrightarrow (x \leqslant y) \wedge (y \leqslant z) \wedge (y \neq z) \Rightarrow (x \leqslant z).$$

现在只需验证 $x \neq z$ 就行了. 如果不然, 就有

$$(x \leqslant y) \wedge (y < z) \Leftrightarrow (z \leqslant y) \wedge (y < z)$$
$$\Leftrightarrow (z \leqslant y) \wedge (y \leqslant z) \wedge (y \neq z).$$

由此据公理 $1_{\leqslant}$ 推知

$$(y = z) \wedge (y \neq z).$$

这是个矛盾. ▶

e. 序与加法及乘法联系的公理的推论 如果与加法、乘法及序公理一起, 还应用联系序与算术运算的公理 (I,III),(II,III), 则可得到下面的断语:

1° *对于任何* $x, y, z, w \in \mathbb{R}$,

$$(x < y) \Rightarrow (x + z) < (y + z),$$
$$(0 < x) \Rightarrow (-x < 0),$$
$$(x \leqslant y) \wedge (z \leqslant w) \Rightarrow (x + z \leqslant y + w),$$
$$(x \leqslant y) \wedge (z < w) \Rightarrow (x + z < y + w).$$

◀ 验证第一个断语.

按严格不等式的定义及公理 (I,III), 得

$$(x < y) \Rightarrow (x \leqslant y) \Rightarrow (x + z) \leqslant (y + z).$$

现在只要验证 $x+z \neq y+z$. 事实上,

$$((x+z)=(y+z)) \Rightarrow (x=(y+z)-z=y+(z-z)=y),$$

这与前提 $x<y$ 矛盾. ►

2° 如果 $x, y, z \in \mathbb{R}$, 那么

$$(0<x) \wedge (0<y) \Rightarrow (0<xy),$$
$$(x<0) \wedge (y<0) \Rightarrow (0<xy),$$
$$(x<0) \wedge (0<y) \Rightarrow (xy<0),$$
$$(x<y) \wedge (0<z) \Rightarrow (xz<yz),$$
$$(x<y) \wedge (z<0) \Rightarrow (yz<xz).$$

◄ 我们来验证第一个断语. 据严格不等式的定义及公理 (II,III)

$$(0<x) \wedge (0<y) \Rightarrow (0 \leqslant x) \wedge (0 \leqslant y) \Rightarrow (0 \leqslant xy).$$

又因前已证明

$$(x \cdot y=0) \Rightarrow (x=0) \vee (y=0),$$

所以 $0 \neq xy$.

再验证第三个断语:

$$\begin{aligned}(x<0) \wedge (0<y) &\Rightarrow (0<-x) \wedge (0<y)\\ &\Rightarrow (0<(-x) \cdot y) \Rightarrow (0<((-1) \cdot x) y)\\ &\Rightarrow (0<-1 \cdot(xy)) \Rightarrow (0<-(xy))\\ &\Rightarrow (xy<0).\end{aligned}$$ ►

读者可自己去证明其余的关系式. 并且同样能验证, 如果上面诸断语左边括号中的一个成为不严格的不等式时, 那么, 右边也要换成不严格的不等式.

3° $0<1$.

◄ $1 \in \mathbb{R} \backslash 0$, 即 $0 \neq 1$.

假如 $1<0$, 那么由刚才证明的结论得知

$$(1<0) \wedge (1<0) \Rightarrow (0<1 \cdot 1) \Rightarrow (0<1).$$

然而我们知道, 任意一对元素 $x, y \in \mathbb{R}$, 在 $x<y, x=y$, 及 $x>y$ 三种情况中, 有且仅有一种可能实现. 因为 $0 \neq 1$, 又假定 $1<0$, 故必导致与它矛盾的关系 $0<1$, 那么, 剩下了一种可能, 这就是断语所说的那种可能. ►

4° $(0 < x) \Rightarrow (0 < x^{-1})$, 且

$$(0 < x) \wedge (x < y) \Rightarrow (0 < y^{-1}) \wedge (y^{-1} < x^{-1}).$$

◀ 先证第一个论断.

首先, $x^{-1} \neq 0$. 假如 $x^{-1} < 0$, 就得

$$(x^{-1} < 0) \wedge (0 < x) \Rightarrow (x \cdot x^{-1} < 0) \Rightarrow (1 < 0).$$

这与前面所证的矛盾. ▶

我们都知道, 大于零的数叫*正数*, 而小于零的数叫*负数*.

因此, 我们已经证明了单位元是正数, 正数与负数的乘积是负数, 正数的倒数也是正数.

3. 完备公理与数集的上 (下) 确界的存在性

定义 2 设 $X \subset \mathbb{R}$ 是一集合; 如果存在一数 $c \in \mathbb{R}$, 使一切 $x \in X$ 都满足 $x \leqslant c (c \leqslant x)$, 就说集合 X 是*上 (下) 有界集*.

这时, 数 c 就叫做 X 的一个*上 (下) 界*.

定义 3 既有上界又有下界的集合叫做*有界集*.

定义 4 元素 $a \in X$ 叫做 $X \subset \mathbb{R}$ 的*最大元或极大元* (*最小元或极小元*), 如果对于一切 $x \in X$ 有 $x \leqslant a$ (或 $a \leqslant x$).

现在引进极大元与极小元定义的表示法, 顺便也就导出它的形式写法:

$$(a = \max X) := (a \in X \wedge \forall x \in X (x \leqslant a)),$$
$$(a = \min X) := (a \in X \wedge \forall x \in X (a \leqslant x)).$$

$\max X$ 读作 "X 的极大元", $\min X$ 读作 "X 的极小元"; 我们也用 $\max\limits_{x \in X} x$ 和 $\min\limits_{x \in X} x$ 来表示 X 的极大元和极小元.

由序公理 $1_{\leqslant}$ 立刻得知, 如一个数集有极大 (极小) 元, 那么它只能有一个.

但是, 并非在任何情况, 甚至在有界集的情况下, 都一定有极大元 (极小元).

例如, 集 $X = \{x \in \mathbb{R} | 0 \leqslant x < 1\}$ 有极小元, 但易知它没有极大元.

定义 5 集合 $X \subset \mathbb{R}$ 的上界中的最小者, 叫做 X 的*上确界*, 写作 $\sup X$ 或 $\sup\limits_{x \in X} x$.

这是本段的基本定义. 于是

$$(s = \sup X) := \forall x \in X ((x \leqslant s) \wedge (\forall s' < s \exists x' \in X (s' < x'))).$$

在上式中, 被定义概念右边的第一个括号说明 s 是 X 的上界, 而第二个括号说明 s 是具有这种性质的数中的最小者. 更确切些说, 第二个括号说明, 比 s 小的数不再是 X 的上界.

类似地可以引进集合 X 的下确界概念, 即集合 X 的下界中的最大者.

定义 6 $(i = \inf X) := \forall x \in X((i \leqslant x) \wedge (\forall i < i' \exists x' \in X(x' < i')))$.
对于集合 X 的下确界, 同记号 $\inf X$ 一样也可以用记号 $\inf\limits_{x\in X} x$.

于是, 我们已经给出了下面的两个定义:

$$\sup X := \min\{c \in \mathbb{R} | \forall x \in X(x \leqslant c)\},$$
$$\inf X := \max\{c \in \mathbb{R} | \forall x \in X(c \leqslant x)\}.$$

然而, 上面我们说过, 并非每个集合一定有极小元或极大元, 所以上述采用的数集的上、下确界的定义需要以下论证.

引理 (上确界原理) 实数集的任何非空有上界的子集有唯一的上确界.

◀ 因为我们已经知道了数集最小元的唯一性, 所以只要证明上确界存在就行了. 设 $X \subset \mathbb{R}$ 是给定的子集, 而 $Y = \{y \in \mathbb{R} | \forall x \in X(x \leqslant y)\}$ 是 X 的上界构成的集合. 由题设 $X \neq \varnothing, Y \neq \varnothing$. 这时根据完备公理, 存在数 $c \in \mathbb{R}$, 使得 $\forall x \in X, \forall y \in Y(x \leqslant c \leqslant y)$. 因此数 c 是 X 的上界, 也是 Y 的下界. 作为 X 的上界, c 是 Y 中的元素, 而作为 Y 的下界, c 是 Y 的极小元. 于是 $c = \min Y = \sup X$. ▶

当然, 类似地可证, 下有界的数集有唯一的下确界, 即

引理 (X 下有界)$\Rightarrow (\exists! \inf X)$.

证明从略.

现在来看集 $X = \{x \in \mathbb{R} | 0 \leqslant x < 1\}$. 由所证的引理, 它应有上确界. 据集 X 的定义及上确界之定义, 显然 $\sup X \leqslant 1$.

这样一来, 为了证明 $\sup X = 1$, 必须验证对任何数 $q < 1$, 能找到数 $x \in X$ 使得 $q < x$, 简单地说就是 q 与 1 之间还有数. 证明这一点当然很容易 (例如可以证明 $q < 2^{-1}(q+1) < 1$), 但是我们现在不打算去证它, 因为在下节中, 我们将对类似的问题逐步而详细地予以讨论.

至于下确界, 则当集合有极小元时, 它的下确界必与集合的极小元一致. 把刚才的这一判断用到我们的例子里, 就得到 $\inf X = 0$.

利用这里所引进概念的内容更丰富的其他例子将在下节中见到.

§2. 最重要的实数类及实数计算方面的一些问题

1. 自然数与数学归纳原理

a. 自然数集的定义 形如 1,1+1,(1+1)+1, 等等的数, 相应地记作 1,2,3, 等等, 叫做*自然数*.

这样一个定义, 只有那样一些人才能接受, 他们在看到这个定义之前, 对自然数(包括它们的记法, 例如十进制记法), 就已经有了完整的概念.

一个过程的延伸, 并非总是一意的, 所以, 到处使用的 "等等", 必须加以明确. 这可用数学归纳基本原理去完成.

定义 1 如果对于集合 $X \subset \mathbb{R}$ 的每个数 $x \in X$, 同时有 $x+1 \in X$, 就称 X 为一个*归纳集*.

例如, $\mathbb{R}$ 是归纳集; 所有正数之集也是归纳集.

任意多个归纳集X_α 之交 $X = \bigcap\limits_{\alpha \in A} X_\alpha$, 如果不空, 也是归纳集.

实际上,

$$\left(x \in X = \bigcap_{\alpha \in A} X_\alpha\right) \Rightarrow (\forall \alpha \in A(x \in X_\alpha))$$
$$\Rightarrow (\forall \alpha \in A((x+1) \in X_\alpha))$$
$$\Rightarrow \left((x+1) \in \bigcap_{\alpha \in A} X_\alpha = X\right).$$

现在引进下面的

定义 2 包含数 1 的最小的归纳集, 即含数 1 的一切归纳集之交, 叫*自然数集*.

自然数集用 $\mathbb{N}$ 表示, 它的元素叫*自然数*.

从集合论的观点看, 可能把自然数集理解成从 0 开始更好些, 即把自然数集作为含有 0 的最小归纳集. 然而对我们来说, 从 1 开始记数更合适.

下面基本而又有广泛应用的原理, 是自然数集定义的直接推论.

b. 数学归纳原理 *如果 E 是自然数集 $\mathbb{N}$ 的子集, $1 \in E$, 并且当 $x \in E$ 时, $x+1$ 也属于 E, 那么 $E = \mathbb{N}$.*

因此,

$$(E \subset \mathbb{N}) \land (1 \in E) \land (\forall x \in E(x \in E \Rightarrow (x+1) \in E)) \Rightarrow E = \mathbb{N}.$$

我们来解释这原理的作用, 借助它证明一些有益的且在以后经常用到的自然数的性质.

1° 自然数的和与积是自然数.

◀ 设 $m,n\in\mathbb{N}$, 今证 $(m+n)\in\mathbb{N}$. 用 E 表示这样的自然数 n 之集, 使对于任何 $m\in\mathbb{N}$ 有 $(m+n)\in\mathbb{N}$. 于是 $1\in E$, 这是因为对于任何 $m\in\mathbb{N}$, $(m\in\mathbb{N})\Rightarrow((m+1)\in\mathbb{N})$. 如果 $n\in E$, 即 $(m+n)\in\mathbb{N}$, 那么由 $(m+(n+1))=((m+n)+1)\in\mathbb{N}$ 知 $(n+1)\in E$. 按归纳原理就得 $E=\mathbb{N}$, 因而证明了加法不会超出 $\mathbb{N}$ 的范围.

类似地, 用 E 表示那些自然数 n 之集, 对任何 $m\in\mathbb{N}$ 有 $(m\cdot n)\in\mathbb{N}$. 于是 $1\in\mathbb{N}$, 这因为 $m\cdot 1=m$; 另外, 如果 $n\in E$, 从而 $m\cdot n\in\mathbb{N}$, 那么 $m\cdot(n+1)=mn+m$ 是二自然数之和, 据刚才已证明了的断言, 它属于 $\mathbb{N}$. 于是 $(n\in E)\Rightarrow((n+1)\in E)$, 据归纳原理, $E=\mathbb{N}$. ▶

2° $(n\in\mathbb{N})\wedge(n\neq 1)\Rightarrow((n-1)\in\mathbb{N})$.

◀ 考察形如 $n-1$ 的自然数之集E, 这里 n 是自然数, 但 n 不是 1. 我们来证 $E=\mathbb{N}$.

因为 $1\in\mathbb{N}$, 那么 $2:=(1+1)\in\mathbb{N}$, 而这表明 $1=2-1\in E$.

如果 $m\in E$, 那么 $m=n-1$, 这里的 $n\in\mathbb{N}$; 这时 $m+1=(n+1)-1$, 由 $(n+1)\in\mathbb{N}$ 有 $m+1\in E$. 据归纳原理推知 $E=\mathbb{N}$. ▶

3° 对于任何 $n\in\mathbb{N}$, 集合 $\{x\in\mathbb{N}|n<x\}$ 有极小元, 并且

$$\min\{x\in\mathbb{N}|n<x\}=n+1.$$

◀ 设 E 为使得上面的断语为真的那些 $n\in\mathbb{N}$ 的集合. 我们来证明 E 与 $\mathbb{N}$ 一致. 首先验证 $1\in E$, 即

$$\min\{x\in\mathbb{N}|1<x\}=2.$$

这个断语也要用归纳原理来验证. 令

$$M=\{x\in\mathbb{N}|(x=1)\vee(2\leqslant x)\},$$

据 M 的定义, $1\in M$. 又如果 $x\in M$, 那么或是 $x=1$, 这时 $x+1=2\in M$; 或是 $2\leqslant x$, 这时 $2\leqslant(x+1)$ 从而也得到 $(x+1)\in M$. 于是 $M=\mathbb{N}$, 这就是说, 如果 $(x\neq 1)\wedge(x\in\mathbb{N})$, 那么 $2\leqslant x$, 这实际上就是 $\min\{x\in\mathbb{N}|1<x\}=2$. 因此, $1\in E$.

现在来证, 如果 $n\in E$, 那么 $n+1\in E$.

首先, 注意 $x\in\{x\in\mathbb{N}|n+1<x\}$, 那么

$$(x-1)=y\in\{y\in\mathbb{N}|n<y\},$$

因为已经证明了所有自然数都不小于 1, 所以 $(n+1<x)\Rightarrow(1<x)\Rightarrow x\neq 1$, 因此, 根据断语 2° 得 $(x-1)=y\in\mathbb{N}$.

现在设 $n\in E$, 亦即 $\min\{y\in\mathbb{N}|n<y\}=n+1$. 这时, $x-1\geqslant y\geqslant n+1$ 且 $x\geqslant n+2$. 这就是说

$$(x\in\{x\in\mathbb{N}|n+1<x\})\Rightarrow(x\geqslant n+2),$$

因之, $\min\{x \in \mathbb{N} | n+1 < x\} = n+2$, 即 $(n+1) \in E$.

根据归纳原理, $E = \mathbb{N}$, 断语 3° 得证. ►

下面的自然数的性质 4°,5°,6°, 都是断语 2°、及 3° 的直接推论:

4° $(m \in \mathbb{N}) \wedge (n \in \mathbb{N}) \wedge (n < m) \Rightarrow (n+1 \leqslant m)$.

5° 数 $(n+1) \in \mathbb{N}$ 是 $\mathbb{N}$ 中紧跟着 n 的一个自然数, 亦即, 当 $n \in \mathbb{N}$ 时, 没有任何自然数 x 能满足条件 $n < x < n+1$.

6° 如果 $n \in \mathbb{N}$, 且 $n \neq 1$, 那么数 $n-1 \in \mathbb{N}$, 并且它是 $\mathbb{N}$ 中在 n 前面的紧挨着 n 的自然数, 即没有任何自然数 x 能满足条件 $n-1 < x < n$.

7° 自然数集的任何非空子集有最小元.

◄ 设 $M \subset \mathbb{N}$. 如果 $1 \in M$, 则因 $\forall n \in \mathbb{N}(1 \leqslant n)$, 所以 $\min M = 1$.

今设 $1 \notin M$, 即 $1 \in E = \mathbb{N} \backslash M$, 在 E 中必能找到这样的自然数 $n \in E$, 使得不超过 n 的自然数都在 E 中, 而 $(n+1) \in M$. 这是因为假如这样的 n 不存在, 那么, 集 $E \subset \mathbb{N}$ 含单位元, 并且当 $n \in E$ 时, E 也包含 $(n+1)$, 从而据归纳原理 E 与 $\mathbb{N}$ 一致. 这不可能, 因为 $\mathbb{N} \backslash E = M \neq \varnothing$.

求得的数 $n+1 \in M$ 必是 M 中的最小元, 因为前已看到, n 与 $n+1$ 之间已不再有自然数了. ►

2. 有理数与无理数

a. 整数

定义 3 自然数集、自然数的相反数之集, 与零的并集, 叫做整数集, 记作 $\mathbb{Z}$.

由于前面已经证明了自然数加法与乘法运算不超出 $\mathbb{N}$ 之范围, 所以, 对整数进行这些运算也不超出 $\mathbb{Z}$ 的范围.

◄ 实际上, 如果 $m, n \in \mathbb{Z}$, 那么或者其中有一个为零, 这时, 和 $m+n$ 就等于另外的那一个, 即 $m+n \in \mathbb{Z}$, 而积 $m \cdot n = 0 \in \mathbb{Z}$; 或者它们都不是零, 这时, 或者 $m, n \in \mathbb{N}$, 从而 $(m+n) \in \mathbb{N} \subset \mathbb{Z}$, 且 $(m \cdot n) \in \mathbb{N} \subset \mathbb{Z}$; 或者 $(-m), (-n) \in \mathbb{N}$, 从而 $m \cdot n = ((-1) \cdot m)((-1) \cdot n) \in \mathbb{N}$; 或者 $(-m), n \in \mathbb{N}$, 这时, $(-m \cdot n) \in \mathbb{N}$, 即 $m \cdot n \in \mathbb{Z}$; 或者, 最后 $m, -n \in \mathbb{N}$, 这时, $(-m \cdot n) \in \mathbb{N}$, 仍有 $m \cdot n \in \mathbb{Z}$. ►

于是, $\mathbb{Z}$ 关于加法运算构成阿贝尔群, 但对乘法运算, 集 $\mathbb{Z}$ 不是群, 甚至 $\mathbb{Z} \backslash 0$ 也不是群, 因为整数的逆 (倒数) 并不属于 $\mathbb{Z}$ (1 与 -1 的倒数除外).

◄ 实际上, 如果 $m \in \mathbb{Z}$ 且 $m \neq 0, 1$, 那么, 先认为 $m \in \mathbb{N}$, 于是 $0 < 1 < m$, 因 $m \cdot m^{-1} = 1 > 0$, 所以应有 $0 < m^{-1} < 1$ (参看上一节中序公理的推论). 因此, $m^{-1} \notin \mathbb{Z}$. 当 m 的负整数且异于 -1 时, 可直接归为已讨论过的情形. ►

当二数 $m, n \in \mathbb{Z}$, 而 $k = m \cdot n^{-1} \in \mathbb{Z}$, 即当 $m = k \cdot n$, 其中 $k \in \mathbb{Z}$ 时, 就说整数 m 能被n 整除, 或它是 n 的倍数, 或说 n 是 m 的因数.

整数的除法, 至多加个负号即乘以 -1 之后, 就变成了相应的自然数的除法, 这些在算术课中都已讨论过了.

我们回顾一下所谓的算术基本定理, 但不给予证明, 在下面讨论某些例子时, 要用到它.

设 $p \in \mathbb{N}, p \neq 1$. 如果在 $\mathbb{N}$ 中除 1 和 p 外, 不再有因数, 我们就把 p 叫做*素数*.

算术基本定理 *每个自然数能唯一地 (不计因数顺序的区别) 表成乘积的形式:*

$$n = p_1 \cdot \cdots \cdot p_k,$$

其中 $p_1, \cdots, p_k$ 都是素数.

数 $m, n \in \mathbb{Z}$ 叫做*互素数*. 假如它们除 $1, -1$ 之外没有公因数.

特别地, 由上述定理能看出, 如果两个互素数 m 与 n 的乘积 $m \cdot n$ 能被素数 p 整除, 那么, 数 m, n 之一也能被 p 整除.

b. 有理数

定义 4 形如 $m \cdot n^{-1}$ 的数叫*有理数*, 其中 $m, n \in \mathbb{Z}$.

有理数集用 $\mathbb{Q}$ 表示[①].

因此, 当 $n \neq 0$ 时, 整数序对 (m, n) 确定了一个有理数.

数 $q = m \cdot n^{-1}$ 也可以写成 m 与 n 之比, 即所谓有理分数 $\frac{m}{n}$ 的形式.

从有理数的定义及实数公理, 马上就能得到在中学学过的与有理数分数表示有关的那些有理数运算规则. 特别有 "将分数的分子与分母乘以同一个不为零的整数, 分数的值不变", 即分数 $\frac{mk}{nk}$ 与 $\frac{m}{n}$ 表示同一个有理数. 实际上, 因为 $(nk)(k^{-1}n^{-1}) = 1$, 即 $(n \cdot k)^{-1} = k^{-1} \cdot n^{-1}$, 所以 $(mk)(nk)^{-1} = (mk)(k^{-1}n^{-1}) = m \cdot n^{-1}$.

因此, 不同的序对 (m, n) 与 (mk, nk) 给出同一个有理数. 因此, 在把任一分数进行相应约简之后, 有理数能够用互素的整数序对给出.

另一方面, 如果序对 (m_1, n_1) 与 (m_2, n_2) 给出同一个有理数, 即 $m_1 \cdot n_1^{-1} = m_2 \cdot n_2^{-1}$, 那么 $m_1 n_2 = m_2 n_1$; 如果 (比如说) m_1 与 n_1 互素, 那么由算术基本定理的推论, 即知 $n_2 \cdot n_1^{-1} = m_2 \cdot m_1^{-1} = k \in \mathbb{Z}$.

这样, 我们证明了两个序对 $(m_1, n_1), (m_2, n_2)$ 给出同一个有理数的充要条件是它们成比例, 即存在数 $k \in \mathbb{Z}$, 使得 (比如说) $m_2 = km_1$ 且 $n_2 = kn_1$.

c. 无理数

定义 5 不是有理数的实数叫*无理数*.

$\sqrt{2}$ 就是这样一个数 $s \in \mathbb{R}, s > 0$ 且 $s^2 = 2$, 它是无理实数的古典例子. 由毕氏定理知, $\sqrt{2}$ 的无理性恰与 "正方形的边与其对角线无公度" 这一命题等价.

---

① $\mathbb{Q}$ 是英语 quotient (商) 的第一个字母.

这样, 我们需验证:第一, 存在正实数 $s\in\mathbb{R}$, 其平方等于 2, 第二, $s\notin\mathbb{Q}$.

◀ 设 X 与 Y 是由这样的正实数构成的集合:$\forall x\in X(x^2<2), \forall y\in Y(2<y^2)$. 因为 $1\in X, 2\in Y$, 所以 X 与 Y 都不空. 又因对于任何正实数 x 与 y 来说,

$$(x<y)\Leftrightarrow(x^2<y^2),$$

所以任何 $x\in X$ 小于任何 $y\in Y$. 据完备公理, 存在着一数 $s\in\mathbb{R}$, 使得 $\forall x\in X, \forall y\in Y$ 有 $x\leqslant s\leqslant y$.

现在来证 $s^2=2$.

假如 $s^2<2$. 那么, $s+\dfrac{2-s^2}{3s}$ 就大于 s, 但它的平方却小于 2. 实际上, 因为 $1\in X$, 所以 $1^2<s^2<2$. 因而 $0<\Delta=2-s^2<1$. 这就是说

$$\begin{aligned}\left(s+\frac{\Delta}{3s}\right)^2&=s^2+2\cdot\frac{\Delta}{3}+\left(\frac{\Delta}{3s}\right)^2<s^2+3\cdot\frac{\Delta}{3}\\&=s^2+\Delta=2.\end{aligned}$$

因此, $\left(s+\dfrac{\Delta}{3s}\right)\in X$, 这与"当 $x\in X$ 时, 必定 $x\leqslant s$"这件事不相容.

假如 $2<s^2$, 那么, $s-\dfrac{s^2-2}{3s}$ 就小于 s, 但它的平方却大于 2. 实际上, 由于 $2\in Y$, 所以 $2<s^2\leqslant 2^2$, 或 $0<\Delta=s^2-2<3$, 从而 $0<\dfrac{\Delta}{3}<1$. 由此推知

$$\begin{aligned}\left(s-\frac{\Delta}{3s}\right)^2&=s^2-2\cdot\frac{\Delta}{3}+\left(\frac{\Delta}{3s}\right)^2>s^2-3\cdot\frac{\Delta}{3}\\&=s^2-\Delta=2.\end{aligned}$$

这又与 s 是集 Y 的下界这个事实矛盾.

于是, 只剩下一种可能, 即 $s^2=2$.

最后证明 $s\notin\mathbb{Q}$. 假如 $s\in\mathbb{Q}$, 设 $\dfrac{m}{n}$ 是 s 的既约分数. 于是 $m^2=2n^2$, 因此 m^2 能被 2 整除, 从而 m 能被 2 整除. 设 $m=2k$, 于是 $2k^2=n^2$. 根据同样的理由, 推知 n 也能被 2 整除, 从而与分数的既约性矛盾. ▶

现在我们致力于证明存在着很多的无理数. 不久就会看到, 在某种意义下, 几乎所有的实数都是无理数. 我们要证明, 无理数集的势, 大于有理数集的势, 而与一切实数之集的势相同.

无理数又可分为所谓代数无理数与超越数.

一个实数, 如果它是某个有理系数 (或等价地说成是整系数的) 代数方程

$$a_0x^n+a_1x^{n-1}+\cdots+a_{n-1}x+a_n=0$$

的根, 就叫做代数数.

反之, 就叫超越数.

我们将看到, 代数数集的势与有理数集的势相同, 而超越数集的势与实数集的势相同. 因此, 乍看起来, 举出一个具体的超越数, 或更确切地说, 证明一个具体的数的超越性也会有很大困难, 似乎不合情理, 也不自然.

然而, 直到 1882 年才证明了著名的几何数 π是超越数①, 而希尔伯特② 问题之一就是证明 α^β 的超越性, 其中的 α 是个代数数 $(\alpha \neq 0) \wedge (\alpha \neq 1)$, 而 β 是代数无理数 (例如, $\alpha = 2, \beta = \sqrt{2}$).

3. 阿基米德③原理 现在来讨论阿基米德原理, 无论从理论方面, 还是从数在测量和计算的具体应用方面看, 它都是重要的. 我们将依据完备公理 (更确切地说, 是依据与之等价的上确界引理) 来证明它. 在其他实数集的公理系统中, 也常把这一基本原理作为公理.

请注意, 到现在为止, 我们已证的有关自然数和整数的定理, 完全没有用到完备公理. 以后将会看到, 阿基米德原理, 实际上反映了与完备公理相联系的自然数和整数的性质. 我们将由此入手.

1° *自然数集的任何不空有界集中有最大元.*

◄ 设 $E \subset \mathbb{N}$ 是任一不空有界集. 根据上确界引理 $\exists! \sup E = s \in \mathbb{R}$. 据上确界之定义, 在 E 中存在自然数 $n \in E$, 满足条件 $s - 1 < n \leqslant s$. 这时 $n = \max E$, 这是因为大于 n 的自然数必不小于 $n + 1$, 而 $n + 1 > s$. ►

推论,2° *自然数集没有上界.*

◄ 如果不然, 那么, 必存在极大的自然数 n. 但 $n < n + 1$. ►

3° *整数集的任何上有界不空子集有极大元.*

◄ 在断语 1° 的证明中, 把 $\mathbb{N}$ 换成 $\mathbb{Z}$, 再重复一遍就行了. ►

4° *自然数集的任何不空有界子集有极小元.*

◄ 比如, 把 $\mathbb{N}$ 换成 $\mathbb{Z}$, 把引用上确界引理改为引用下确界引理, 再重复断语 1° 的证明就行了.

也可以变成相反数 (“改变符号”) 并利用 3° 中已证明的结果. ►

5° *整数集既没有上界又没有下界.*

◄ 由 3°,4°, 或直接由 2° 推知. ►

现在叙述

---

① π 是一个数, 它等于欧氏几何中圆周长与直径之比. 因此, 从 18 世纪起人们就采用 περιφερια (希腊文, 圆周之意) 的第一个字母来表示这个数. π 的超越性是由德国数学家林达曼 (Lindemann) 证明的. 特别地, 由 π 的超越性可得出用直尺和圆规作长度为 π 的线段 (圆周拉直的问题) 是不可能的, 同样地, 用这些工具化圆为方 (这是中世纪的一个问题) 也是不可能的.

② D. Hilbert (1862—1943)—— 卓越的德国数学家, 在 1900 年巴黎数学国际会议上提出了 23 个涉及数学各个领域的问题, 这些问题后来就叫做 “Hilbert 问题”. 本文中所提到的问题 (Hilbert 第七问题) 在 1934 年被苏联数学家盖尔芳德 (А. О. Гельфонд) (1906—1968) 以肯定的结论解决了.

③ Archimedes (公元前 287—前 212)—— 天才的希腊学者. 分析奠基人之一的莱布尼茨, 在谈到他时说: “研究了阿基米德的著作后, 就不再对现代数学的成就感到惊奇了”.

6° 阿基米德原理 如果 h 是任意一个固定的正数, 那么, 对于任何实数 x, 必能找到唯一的整数 k, 使得 $(k-1)h \leqslant x < kh$.

◀ 因为 $\mathbb{Z}$ 没有上界, 集合 $\left\{n \in \mathbb{Z} \middle| \frac{x}{h} < n\right\}$ 是整数集的一个不空下有界子集. 这时 (由 4°) 其中有最小元 k, 即 $(k-1) \leqslant \frac{x}{h} < k$. 因为 $h > 0$, 这些不等式与阿基米德原理中所说的不等式等价. 满足上面两个不等式的整数 k 是唯一的, 由数集的极小元的唯一性推知 (参看 §1, 第 3 小段). ▶

几个推论:

7° 对于任意的正数 ε, 存在着自然数 n 使得 $0 < \frac{1}{n} < \varepsilon$.

◀ 据阿基米德原理, 存在 $n \in \mathbb{Z}$ 使得 $1 < \varepsilon \cdot n$. 因为 $0 < 1$, 且 $0 < \varepsilon$, 所以 $0 < n$. 于是, $n \in \mathbb{N}$, 且 $0 < \frac{1}{n} < \varepsilon$. ▶

8° 如果 $x \in \mathbb{R}, x \geqslant 0$, 且对于任何 $n \in \mathbb{N}$ 有 $x < \frac{1}{n}$, 那么 $x = 0$.

◀ 据断语 7°, 不可能有 $x > 0$. ▶

9° 对于满足 $a < b$ 的任意数 $a, b \in \mathbb{R}$, 存在有理数 $r \in \mathbb{Q}$, 使得 $a < r < b$.

◀ 回顾 7°, 可选出某个 $n \in \mathbb{N}$, 使得 $0 < \frac{1}{n} < b - a$. 据阿基米德原理, 存在着数 $m \in \mathbb{Z}$, 使得 $\frac{m-1}{n} \leqslant a < \frac{m}{n}$. 这时, $\frac{m}{n} < b$, 这是因为, 若不然, 就有 $\frac{m-1}{n} \leqslant a < b < \frac{m}{n}$, 由此得 $\frac{1}{n} > b - a$. 因此, $r = \frac{m}{n} \in \mathbb{Q}$ 且 $a < \frac{m}{n} < b$. ▶

10° 对于任何 $x \in \mathbb{R}$, 存在唯一的整数 $k \in \mathbb{Z}$, 使得 $k \leqslant x < k + 1$.

◀ 由阿基米德原理直接推知. ▶

10° 中的这个数 k 用 $[x]$ 来记, 叫做 x 的整数部分. 量 $\{x\} := x - [x]$ 叫做 x 的小数部分. 因此, $x = [x] + \{x\}$ 并且 $\{x\} \geqslant 0$.

4. 实数集的几何解释与实数计算方面的一些问题

a. 数轴 对于实数经常使用形象的几何语言, 这在中学里就大体上已经知道了. 根据几何公理, 直线 $\mathbb{L}$ 上的点与实数集 $\mathbb{R}$ 的数之间能建立一一对应 $f : \mathbb{L} \to \mathbb{R}$, 并且这个对应与直线的运动有关, 如果 T 是直线 $\mathbb{L}$ 沿自身的一个平行位移, 那么存在一个 (只与 T 有关的) 数 $t \in \mathbb{R}$, 使得对于任何点 $x \in \mathbb{L}$ 有 $f(T(x)) = f(x) + t$.

与点 $x \in \mathbb{L}$ 相对应的数 $f(x)$ 叫做点 x 的坐标. 由于映射 $f : \mathbb{L} \to \mathbb{R}$ 是一一的, 所以也常直接把点的坐标叫做点, 例如把 "我们标出坐标为 1 的点" 说成 "我们标出点 1". 当存在上面所说的对应 $f : \mathbb{L} \to \mathbb{R}$ 时, 称直线为*坐标轴*, *数轴*或*数直线*. 由于 f 是双射, 实数集 $\mathbb{R}$ 本身也常叫做数直线, 而它的元素, 叫做数直线上的点.

我们已经看到, 给出 $\mathbb{L}$ 上坐标的双射 $f : \mathbb{L} \to \mathbb{R}$, 当平移 T 时, 直线 $\mathbb{L}$ 上的点的像的坐标与该点的坐标相差的是同一个量 $t \in \mathbb{R}$. 因此, 一旦坐标为 1 的点和坐标为 0 的点被确定, 映射 f 就完全确定了, 简单地说, f 被零点 (称为*坐标原点*) 及点 1 完全确定. 由这二点所确定的线段叫*单位线段*. 从原点出发而含有点 1 的射线所定的

方向叫正方向, 而沿这个方向 (从 0 向 1) 的运动叫做自左向右的运动, 按照这种约定, 就说 1 位于 0 的右边, 而 0 位于 1 的左边.

在坐标原点 x_0 变到坐标为 1 的点 $x_1 = T(x_0)$ 的这个平移下, 所有点的像的坐标都较原像的坐标增大了 1. 因此, 我们发现坐标为 2 的点是 $x_2 = T(x_1)$, 坐标为 3 的点是 $x_3 = T(x_2), \cdots$, 坐标为 $n+1$ 的点是 $x_{n+1} = T(x_n)$, 以及坐标为 -1 的点是 $x_{-1} = T^{-1}(x_0), \cdots$ 坐标为 $-n-1$ 的点是 $x_{-n-1} = T^{-1}(x_n)$. 于是, 我们得到了坐标为整数 $m \in \mathbb{Z}$ 的一切点.

我们会把单位线段增至二倍, 三倍, $\cdots$, 根据泰勒斯 (Thales) 定理也就能把单位线段分成相应的 n 个合同的线段, 取其中有一个端点为原点的线段, 则另一端的坐标 x 满足等式 $n \cdot x = 1$, 即 $x = \frac{1}{n}$. 由此即可求得具有有理坐标 $\frac{m}{n} \in \mathbb{Q}$ 的一切点.

然而在 $\mathbf{L}$ 上还有其他的点, 因为有的线段是与单位线段无公度的. 每个这样的点 (就像直线上的其他点一样) 把直线分成了两个射线, 在每个射线上有一些以整数 (有理数) 为坐标的点 (这是阿基米德原来的几何原理的推论). 这样一来, 点产生出了分割, 或如通常所说, 产生出了 $\mathbb{Q}$ 的分划, 把 $\mathbb{Q}$ 分成两个非空集 X 与 Y, 分别对应着左、右射线上的有理点 (以有理数为坐标的点). 据完备公理, 存在将 X 和 Y 分开的数 c, 即 $\forall x \in X$ 及 $\forall y \in Y$ 有 $x \leqslant c \leqslant y$. 因为 $X \cup Y = \mathbb{Q}$, 所以 $\sup X = s = i = \inf Y$, 这是由于若不如此, 就有 $s < i$, 从而在 s 与 i 之间就找到一个既不在 X 中又不在 Y 中的有理数. 因此, $s = i = c$. 这个唯一确定的数 c 就对应于直线上指定的点.

上述直线上的点与其坐标的对应, 给出了 $\mathbb{R}$ 的一个模型, 无论从 $\mathbb{R}$ 的序关系方面 (据此得到 "线性序" 这个名称), 还是从 $\mathbb{R}$ 的完备 (连续) 公理方面看, 这个模型都是很直观的. $\mathbb{R}$ 的完备 (连续) 公理用几何语言说就是, 直线 $\mathbf{L}$ 上 "没有洞" 能把 $\mathbf{L}$ 分成没有公共点的两块 (这样的分划只能用直线自己的点实现).

下面, 我们不去深入讨论映射 $f : \mathbf{L} \to \mathbb{R}$ 的结构的细节, 因为我们介绍实数集的几何解释只是为了更加直观, 更能激发读者的极为有益的几何直觉. 至于形式化证明, 仍像往常一样, 或者根据一组已经从公理得到的事实, 或者根据这个公理系统进行.

往后我们要经常使用几何语言.

今对以下所列出的集合引进记号和名称:

$$]a, b[:= \{x \in \mathbb{R} | a < x < b\} \text{—— 开区间} ab;$$

$$[a, b] := \{x \in \mathbb{R} | a \leqslant x \leqslant b\} \text{—— 闭区间} ab;$$

$$]a, b] := \{x \in \mathbb{R} | a < x \leqslant b\} \text{—— 含端点} b \text{的半开区间 } ab;$$

$$[a, b[:= \{x \in \mathbb{R} | a \leqslant x < b\} \text{—— 含端点} a \text{的半开区间 } ab.$$

定义 6 开区间、闭区间、半开区间都叫数区间或简称区间, 确定区间的数叫做它的端点.

量 $b-a$ 叫做区间 ab 的长. 如果 I 是某个区间, 那么, 就用 $|I|$ 表示它的长 (这个记法的来由, 很快就会知道).

集合

$$\begin{aligned}
]a,+\infty[&:= \{x \in \mathbb{R} | a < x\}, \\
[a,+\infty[&:= \{x \in \mathbb{R} | a \leqslant x\}, \\
]-\infty,b[&:= \{x \in \mathbb{R} | x < b\}, \\
]-\infty,b] &:= \{x \in \mathbb{R} | x \leqslant b\}, \\
]-\infty,+\infty[&:= \mathbb{R},
\end{aligned}$$

都是无界集叫*无界区间*.

与符号 $+\infty$ (读作 "正无穷") 及 $-\infty$ (读作 "负无穷") 的用法相适应, 为了表达数集 X 的上 (下) 无界性, 也采用 $\sup X = +\infty(\inf X = -\infty)$ 这种写法.

定义 7　称含有 $x \in \mathbb{R}$ 的开区间为 x 的*一个邻域*.

特别当 $\delta > 0$ 时, 开区间 $]x-\delta, x+\delta[$ 叫做点 x 的*一个 δ 邻域*, 它的长是 2δ.

数 $x, y \in \mathbb{R}$ 之间的距离, 就是以 x, y 为端点的区间之长.

为了不必区别 "哪点在左, 哪点在右", 即 $x < y$ 或 $y < x$, 其长度等于 $y-x$ 或 $x-y$, 可以使用下面这个很方便的函数

$$|x| = \begin{cases} x, & \text{当}x > 0, \\ 0, & \text{当}x = 0, \\ -x, & \text{当}x < 0, \end{cases}$$

这函数称为数 x 的*模*或*绝对值*.

定义 8　称 $|x-y|$ 为 $x, y \in \mathbb{R}$ 之间的*距离*.

距离是非负的, 而且, 当且仅当 x 与 y 重合时, 它才等于零. 从 x 到 y 的距离与从 y 到 x 的距离是一样的, 这是因为 $|x-y| = |y-x|$. 最后, 如果 $z \in \mathbb{R}$, 那么

$$|x-y| \leqslant |x-z| + |z-y|,$$

即所谓三角形不等式成立.

三角形不等式可由数的绝对值的性质推出来, 这个性质也叫三角形不等式 (在上面那个不等式中, 令 $z=0$, 及以 $-y$ 代 y 即可得到), 即对于任何数 x, y, 下面不等式成立:

$$|x+y| \leqslant |x| + |y|,$$

并且等式成立的条件是 x 与 y 同时非负或同时非正.

◀ 如果 $0 \leqslant x$ 且 $0 \leqslant y$, 那么 $0 \leqslant x+y, |x+y| = x+y, |x| = x, |y| = y$, 这时等式成立.

如果 $x \leqslant 0, y \leqslant 0$, 那么 $x+y \leqslant 0, |x+y| = -(x+y) = -x-y, |x| = -x, |y| = -y$, 这时也是等式成立.

今设 x, y 二数中有一个是负的, 另一个是正的, 例如, $x < 0 < y$. 这时, 或者 $x < x+y \leqslant 0$; 或者 $0 \leqslant x+y \leqslant y$. 第一种情况时 $|x+y| < |x|$, 第二种情况时, $|x+y| < |y|$, 从而, 不管在哪种情况都有 $|x+y| < |x| + |y|$. ▶

利用归纳原理可以验证

$$|x_1 + \cdots + x_n| \leqslant |x_1| + \cdots + |x_n|,$$

并且等式成立的充要条件是所有数 $x_1, \cdots, x_n$ 同时非负或同时非正.

数 $\dfrac{a+b}{2}$ 叫做 a, b 的中点, 或叫做以 a, b 为端点的区间的中心, 因为它距区间的两端点等远.

特别地, 点 $x \in \mathbb{R}$ 是它的 δ 邻域 $]x-\delta, x+\delta[$ 的中心, 而 δ 邻域中的一切点与 x 的距离小于 δ.

b. 用逼近序列给出数 测量一个实际的物理量就得到一个数, 通常在重复进行测量时, 这个测得的数会有变动, 当测量工具和测量方法改变时尤其如此. 因此, 测量的结果通常是所求量的一个近似值. 测量的好坏或精确度, 可以用 (例如) 量的真正的值与测得的值的可能有的偏差量来刻画. 同时, 可能出现总也测不到该量精确值 (如果它原则上是*存在*的话) 的情况. 然而, 从更加构造性的观点出发, 如果我们有能力以任意的精确度进行测量的话, 就可以 (或应当) 认为完全知道所求的量. 这种观点就是把一个数与它的经测量得到的越来越精确的近似值序列看成是同一的[①]. 但是, 为了测量, 首先要选定一个标准尺度. 我们称标准尺度的与整个标准尺度可公度的部分为标准尺度的有理部分. 易见, 任何测量都是与标准尺度, 或与标准尺度的有理部分所作的有限次比较, 因此, 测量的结果用自然数, 整数, 一般地, 用有理数来表示. 这就是说, 经过必须的分析, 建立了实数集的数学拷贝, 说得更明确些, 是建立了那样一个模型, 它使我们在处理实数时丝毫也不怀疑它们的公理化描述. 此后, 原则上就能用有理数列来描述整个实数集. 而待测的未知数的加法与乘法, 则用它们的近似值来代替.(的确, 并非总能说清楚, 这样运算的结果与将这些运算施于精确值时所得的结果之间有什么关系. 下面我们将讨论这一问题.)

在把数等同于它的近似值序列后, 为 (比如) 把两数相加, 应当去加它们的近似值序列, 这时必须把所得到的新序列看成一个新数, 它叫做前面那两数之和. 然而这是不是一个数呢? 微妙之处在于, 并非每个随便构造的序列都能作为某个量的任意精确的近似值序列. 因此, 还需研究如何根据序列本身来判断它是不是一个数. 在数

① 若 n 是测量的序号, 而 x_n 是测量结果, 那么对应 $n \mapsto x_n$ 按序列之定义, 不过是自然数变量的一个函数 $f: \mathbb{N} \to \mathbb{R}$ (在所给的情况是数列) 在第三章 §1 中将对数列作详细的讨论

学模拟近似数运算时, 产生的另一个问题是, 不同序列可能是同一个量的近似序列. 确定一个数的那些近似值序列与数本身之间的关系, 大体就像图上的点与指示该点的教鞭之间的关系一样. 教鞭指出了点的位置, 而点确定的只是教鞭端点的位置, 它不会影响你使用另外一个更方便的教鞭.

对这些问题, 柯西[①]作了精确的描述, 并大体上实现了这里所提出的建立实数模型的纲要. 可以期望, 在学习了极限理论之后, 你们将能独立地重复出柯西的这些结构.

当然, 到现在为止, 所说的一切并没追求数学的严格性. 这些看似离题的议论的目的是, 提醒读者注意, 原则上可同时存在实数的各种自然科学模型; 我还试图给出有关数与我们周围事物关系的一些概念, 并阐明自然数与有理数的基本作用; 最后, 想解释一下近似计算的必然性与必要性.

本段后半部分讲述有关近似值算术运算的简单而重要的误差估计问题, 这些知识今后将会用到, 也有自己独立的趣味.

现在转入对这些问题的确切叙述.

定义 9 设 x 是某个量的精确值, $\bar{x}$ 是该量的已知近似值, 就把

$$\Delta(\bar{x}) := |x - \bar{x}|, \quad \delta(\bar{x}) := \frac{\Delta(\bar{x})}{|\bar{x}|}$$

分别叫做近似值 $\bar{x}$ 的*绝对误差*与*相对误差*. 当 $\bar{x} = 0$ 时, 相对误差没有定义.

因为值 x 未知, 所以 $\Delta(\bar{x})$ 与 $\delta(\bar{x})$ 也不知道. 然而通常这些量的上估计 $\Delta(\bar{x}) < \Delta, \delta(\bar{x}) < \delta$, 是知道的. 这时我们说近似值 $\bar{x}$ 的绝对误差与相对误差分别不超过 Δ 与 δ. 实践中常用的只是误差的估值, 所以常常把 Δ 与 δ 叫做近似值的*绝对误差*与*相对误差*, 但是我们不这样称呼它们.

记法 $x = \bar{x} \pm \Delta$ 表示 $\bar{x} - \Delta \leqslant x \leqslant \bar{x} + \Delta$.

例如,

万有引力常数	$G = (6.672\,59 \pm 0.000\,85) \cdot 10^{-11}$ 牛顿 · 米^2/千克^2,
真空中的光速	$c = 299\,792\,458$ 米/秒 (精确),
普朗克常数	$h = (6.626\,075\,5 \pm 0.000\,004\,0) \cdot 10^{-34}$ 焦耳 · 秒,
电子的 (基本) 电荷	$e = (1.602\,177\,33 \pm 0.000\,000\,49) \cdot 10^{-19}$ 库仑,
电子的 (静) 质量	$m_e = (9.109\,389\,7 \pm 0.000\,005\,4) \cdot 10^{-31}$ 千克.

近似值的相对误差是测量精确度的一个基本标志, 通常用百分数表示它.

例如, 前面诸例中的相对误差分别不超过

$$13 \cdot 10^{-5}; \quad 0; \quad 6 \cdot 10^{-7}; \quad 31 \cdot 10^{-8}; \quad 6 \cdot 10^{-7}$$

---

① 柯西(Cauchy)(1789—1857)—— 法国数学家, 分析的现代语言与机器的最积极的创作者之一.

或写成百分数时, 就是

$$13\cdot 10^{-3}\%,\quad 0\%;\quad 6\cdot 10^{-5}\%;\quad 31\cdot 10^{-6}\%;\quad 6\cdot 10^{-5}\%.$$

现在估计用近似值做算术运算时产生的误差.

命题 如果

$$|x-\tilde{x}|=\Delta(\tilde{x}),\quad |y-\tilde{y}|=\Delta(\tilde{y}),$$

那么

$$\Delta(\tilde{x}+\tilde{y}):=|(x+y)-(\tilde{x}+\tilde{y})|\leqslant \Delta(\tilde{x})+\Delta(\tilde{y}),\tag{1}$$

$$\Delta(\tilde{x}\cdot\tilde{y}):=|x\cdot y-\tilde{x}\cdot\tilde{y}|\leqslant |\tilde{x}|\Delta(\tilde{y})+|\tilde{y}|\Delta(\tilde{x})+\Delta(\tilde{x})\cdot\Delta(\tilde{y});\tag{2}$$

如果再设

$$y\neq 0,\tilde{y}\neq 0\quad \text{且}\quad \delta(\tilde{y})=\frac{\Delta(\tilde{y})}{|\tilde{y}|}<1,$$

那么

$$\Delta\left(\frac{\tilde{x}}{\tilde{y}}\right):=\left|\frac{x}{y}-\frac{\tilde{x}}{\tilde{y}}\right|\leqslant\frac{|\tilde{x}|\Delta(\tilde{y})+|\tilde{y}|\Delta(\tilde{x})}{\tilde{y}^2}\cdot\frac{1}{1-\delta(\tilde{y})}.\tag{3}$$

◀ 令 $x=\tilde{x}+\alpha, y=\tilde{y}+\beta$, 则

$$\begin{aligned}
\Delta(\tilde{x}+\tilde{y})&=|(x+y)-(\tilde{x}+\tilde{y})|=|\alpha+\beta|\\
&\leqslant|\alpha|+|\beta|=\Delta(\tilde{x})+\Delta(\tilde{y}),\\
\Delta(\tilde{x}\cdot\tilde{y})&=|xy-\tilde{x}\tilde{y}|=|(\tilde{x}+\alpha)(\tilde{y}+\beta)-\tilde{x}\tilde{y}|\\
&=|\tilde{x}\beta+\tilde{y}\alpha+\alpha\beta|\leqslant|\tilde{x}||\beta|+|\tilde{y}||\alpha|+|\alpha\beta|\\
&=|\tilde{x}|\Delta(\tilde{y})+|\tilde{y}|\Delta(\tilde{x})+\Delta(\tilde{x})\cdot\Delta(\tilde{y}),\\
\Delta\left(\frac{\tilde{x}}{\tilde{y}}\right)&=\left|\frac{x}{y}-\frac{\tilde{x}}{\tilde{y}}\right|=\left|\frac{x\tilde{y}-y\tilde{x}}{y\tilde{y}}\right|\\
&=\left|\frac{(\tilde{x}+\alpha)\tilde{y}-(\tilde{y}+\beta)\tilde{x}}{\tilde{y}^2}\right|\left|\frac{1}{1+\beta/\tilde{y}}\right|\\
&\leqslant\frac{|\tilde{x}||\beta|+|\tilde{y}||\alpha|}{\tilde{y}^2}\cdot\frac{1}{1-\delta(\tilde{y})}\\
&=\frac{|\tilde{x}|\Delta(\tilde{y})+|\tilde{y}|\Delta(\tilde{x})}{\tilde{y}^2}\cdot\frac{1}{1-\delta(\tilde{y})}.
\end{aligned}$$ ▶

由所得到的绝对误差的估值, 可得相对误差的以下估计:

$$\delta(\tilde{x}+\tilde{y})\leqslant\frac{\Delta(\tilde{x})+\Delta(\tilde{y})}{|\tilde{x}+\tilde{y}|},\tag{1'}$$

$$\delta(\tilde{x}\cdot\tilde{y})\leqslant\delta(\tilde{x})+\delta(\tilde{y})+\delta(\tilde{x})\cdot\delta(\tilde{y}),\tag{2'}$$

$$\delta\left(\frac{\tilde{x}}{\tilde{y}}\right) \leqslant \frac{\delta(\tilde{x})+\delta(\tilde{y})}{1-\delta(\tilde{y})}. \tag{3'}$$

在实践中, 当近似值足够好时, 有 $\Delta(\tilde{x})\cdot\Delta(\tilde{y})\approx 0, \delta(\tilde{x})\cdot\delta(\tilde{y})\approx 0, 1-\delta(\tilde{y})\approx 1$, 所以得到 (2),(3),(2′),(3′) 的有用的简化形式:

$$\begin{aligned}
&\Delta(\tilde{x}\cdot\tilde{y}) \leqslant |\tilde{x}|\Delta(\tilde{y})+|\tilde{y}|\Delta(\tilde{x}),\\
&\Delta\left(\frac{\tilde{x}}{\tilde{y}}\right) \leqslant \frac{|\tilde{x}|\Delta(\tilde{y})+|\tilde{y}|\Delta(\tilde{x})}{\tilde{y}^2},\\
&\delta(\tilde{x}\cdot\tilde{y}) \leqslant \delta(\tilde{x})+\delta(\tilde{y}),\\
&\delta\left(\frac{\tilde{x}}{\tilde{y}}\right) \leqslant \delta(\tilde{x})+\delta(\tilde{y}).
\end{aligned}$$

但这些公式并不是由 (2),(3),(2′),(3′) 精确地转换来的.

公式 (3),(3′) 说明, 当 $\tilde{y}$ 或 $1-\delta(\tilde{y})$ 的绝对值很小时, 必须避免用近于零或十分粗糙的近似值做除法的除数.

公式 (1′) 警告我们, 在做近似值的加法时, 如果两个数的绝对值接近但符号相反, 就要特别当心, 因为这时 $|\tilde{x}+\tilde{y}|$ 近于零.

在所有这些情形下, 误差可能增大得很厉害.

例如, 用某种仪器两次测量你的身高. 设测量的精确度为 ± 0.5 cm, 第二次测量前, 在你两脚下先垫上一张薄纸. 虽然如此, 两次测量结果仍可能是 $H_1=(200\pm 0.5)$cm 与 $H_2=(199.8\pm 0.5)$cm.

这样, 借两次测量之差 H_2-H_1 以求纸的厚度就是无意义的事了, 因为, 从这里只能得出厚度不超过 0.8 cm. 当然, 对纸的厚度的这样一种反映 (如果也算是一种"反映" 的话), 就显得太粗糙了.

然而值得注意的是另外的比较乐观的计算效应, 由于这种效应, 用粗糙工具也可以进行精细的测量. 例如, 还用那件刚才为你测身高的仪器, 测了一叠 1 000 页同样的纸的厚度, 测得的结果是 (20 ± 0.5)cm. 那么, 据公式 (3) 得知* 一张纸的厚度就是$(0.02\pm 0.000\,5)$cm $=(0.2\pm 0.005)$mm.

这就是说一张纸的厚度等于 0.2 mm, 绝对误差不超过 0.005 mm, 相对误差不超过 0.025 或 2.5%.

c. 位置记数法 上面说到每个数可用它的有理近似序列给出来.

现在介绍一种在计算上很重要的方法, 对于每个实数, 能建立唯一的这样的有理近似数列. 这种方法导致了位置记数法的建立.

引理 *如果固定数 $q>1$, 那么, 对于任何正数 $x\in\mathbb{R}$, 必有唯一的整数 $k\in\mathbb{Z}$, 使得*

$$q^{k-1}\leqslant x<q^k.$$

*译者注 原文为根据公式 (1)

◀ 首先证明形如 $q^k(k\in\mathbb{N})$ 的数没有上界. 如若不然, 它就有上确界 s, 从而, 据上确界之定义, 能求出自然数 $m\in\mathbb{N}$, 使得 $\frac{s}{q}<q^m\leqslant s$. 这时 $s<q^{m+1}$, 因而 s 就不能是我们这个集的上确界了.

由于 $q>1$, 得知当 $m<n$, 且 $m,n\in\mathbb{Z}$ 时, 有 $q^m<q^n$, 因此, 我们一并证明了, 对于任何 $c\in\mathbb{R}$, 能找到自然数 $N\in\mathbb{N}$, 使得对任何自然数 $n>N$, 必有 $c<q^n$.

由此推知, 对任何 $\varepsilon>0$, 能求得 $M\in\mathbb{N}$, 使得对于任何自然数 $m>M$, 有 $\frac{1}{q^m}<\varepsilon$.

实际上, 只要令 $c=\frac{1}{\varepsilon}, N=M$, 那么, 当 $m>M$ 时, 就有 $\frac{1}{\varepsilon}<q^m$.

于是, 对 $x>0$, 满足不等式 $x<q^m$ 的整数 $m\in\mathbb{Z}$ 的集合有下界. 这时, 这集合有最小元 k. 显然, 这个 k 就是所求的整数, 因为对于它成立 $q^{k-1}\leqslant x<q^k$.

这样的整数 k 的唯一性, 可由以下事实推知: 如果 $m,n\in\mathbb{Z}$, 比如说 $m<n$, 那么 $m\leqslant n-1$, 从而当 $q>1$ 时 $q^m\leqslant q^{n-1}$.

实际上, 因为由不等式 $q^{m-1}\leqslant x<q^m$ 及 $q^{n-1}\leqslant x<q^n$ 可得 $q^{n-1}\leqslant x<q^m$, 因此, 由上面这个注记看到, 如果 $m\neq n$, 它们必不能同时成立. ▶

在下面的结构中, 要用到这个引理.

固定 $q>1$, 取任意的正数 $x\in\mathbb{R}$.

根据引理可求得唯一的整数 $p\in\mathbb{Z}$, 使得

$$q^p\leqslant x<q^{p+1}. \tag{1}$$

定义 10 满足关系 (1) 的整数 p 叫做数 x关于记数法的基 q 的阶或 (当把 q 固定时) 简称为数 x 的阶.

根据阿基米德原理, 存在唯一的自然数 $\alpha_p\in\mathbb{N}$, 使得

$$\alpha_pq^p\leqslant x<\alpha_pq^p+q^p. \tag{2}$$

注意到 (1) 式, 即可断定 $\alpha_p\in\{1,\cdots,q-1\}$.

在我们的构造中, 后继的每一步, 都是重复马上进行的从 (2) 式出发的如下步骤.

由关系式 (2) 及阿基米德原理推知, 存在着唯一的数 $\alpha_{p-1}\in\{0,1,\cdots,q-1\}$, 使得

$$\alpha_pq^p+\alpha_{p-1}q^{p-1}\leqslant x<\alpha_pq^p+\alpha_{p-1}q^{p-1}+q^{p-1}. \tag{3}$$

假如已经做完这样的 n 步, 并得到

$$\alpha_pq^p+\alpha_{p-1}q^{p-1}+\cdots+\alpha_{p-n}q^{p-n}\leqslant x<\alpha_pq^p+\alpha_{p-1}q^{p-1}+\cdots+\alpha_{p-n}q^{p-n}+q^{p-n},$$

那么, 根据阿基米德原理能求得唯一的数$\alpha_{p-n-1} \in \{0,1,\cdots,q-1\}$使得

$$\alpha_p q^p + \cdots + \alpha_{p-n} q^{p-n} + \alpha_{p-n-1} q^{p-n-1} \leqslant x < \alpha_p q^p + \cdots + \alpha_{p-n} q^{p-n} + \cdots + \alpha_{p-n-1} q^{p-n-1} + q^{p-n-1}.$$

于是制定了一种算法, 根据它, 每个正数 x 一意地与数列 $\alpha_p, \alpha_{p-1}, \cdots, \alpha_{p-n}, \cdots$ 对应, 其中诸 α_j 都是 $\{0,1,\cdots,q-1\}$ 中的数. 或者用不太形式化的说法, 每个数 x 一意地与具有特别形式的有理数

$$r_n = \alpha_p q^p + \cdots + \alpha_{p-n} q^{p-n} \tag{4}$$

的序列对应, 这里

$$r_n \leqslant x < r_n + \frac{1}{q^{n-p}}. \tag{5}$$

换句话说, 我们用特殊的有理数序列 (4), 建立了 x 的越来越好的不足近似值与过剩近似值. 符号 $\alpha_p, \cdots, \alpha_{p-n}, \cdots$ 是整个序列 $\{r_n\}$ 的密码, 为了能根据这个密码恢复序列 $\{r_n\}$, 必须指出 x 的阶数 p.

我们约定, 当 $p \geqslant 0$ 时, 在 α_0 后面放上小数点或逗点; 当 $p < 0$ 时, 在 α_p 的左边补写上 $|p|$ 个零, 而在最左边的零后边放上小数点或逗点 (注意 $\alpha_p \neq 0$).

例如, 当 $q = 10$ 时,

$$123.45 := 1 \cdot 10^2 + 2 \cdot 10^1 + 3 \cdot 10^0 + 4 \cdot 10^{-1} + 5 \cdot 10^{-2},$$
$$0.001\,23 := 1 \cdot 10^{-3} + 2 \cdot 10^{-4} + 3 \cdot 10^{-5};$$

当 $q = 2$ 时,

$$1\,000.001 := 1 \cdot 2^3 + 1 \cdot 2^{-3}.$$

因此, 符号 $\alpha_p, \cdots, \alpha_{p-n}, \cdots$ 中数码的值, 依赖于它相对于小数点或逗点的位置.

在做了这些约定之后, 就可从符号 $\alpha_p, \cdots, \alpha_0, \cdots$ 一意地恢复整个的近似序列来.

由 (5) 式看出 (试验证!), 两个不同的数 x 与 x' 对应着不同的序列, 即不同的符号 $\alpha_p, \cdots, \alpha_0, \cdots$ 与 $\alpha'_p, \cdots, \alpha'_0, \cdots$.

现在问, 是否形如与 $\alpha_p, \cdots, \alpha_0, \cdots$ 的每个符号对应都有一个数呢? 答案竟然是否定的.

请注意, 根据上述逐次得到 $\alpha_{p-n} \in \{0,1,\cdots,q-1\}$ 的算法, 不会发生从某项开始以后所有的 α_{p-n} 都等于 $q-1$ 的情形.

实际上, 如果当 $n > k$ 时, 有

$$r_n = \alpha_p q^p + \cdots + \alpha_{p-k} q^{p-k} + (q-1) q^{p-(k+1)} + \cdots + (q-1) q^{p-n},$$

即

$$r_n = r_k + \frac{1}{q^{k-p}} - \frac{1}{q^{n-p}}, \tag{6}$$

那么, 由 (5) 式得

$$r_k + \frac{1}{q^{k-p}} - \frac{1}{q^{n-p}} \leqslant x < r_k + \frac{1}{q^{k-p}},$$

这时, 对于任何 $n > k$

$$0 < r_k + \frac{1}{q^{k-p}} - x < \frac{1}{q^{n-p}},$$

而由上所证之引理, 这是不可能的.

还应注意, 如果在 $\alpha_{p-k-1}, \cdots, \alpha_{p-n}$ 中即使有一个小于 $q-1$, 那么代替 (6) 式有

$$r_n < r_k + \frac{1}{q^{k-p}} - \frac{1}{q^{n-p}},$$

亦即

$$r_n + \frac{1}{q^{n-p}} < r_k + \frac{1}{q^{k-p}}. \tag{7}$$

现在我们可以证明, 由数 $\alpha_k \in \{0, 1, \cdots, q-1\}$ 组成的任何记号 $\alpha_p, \cdots, \alpha_0, \cdots$, 如果其中无论哪一个数码后面总有不是 $q-1$ 的, 则它必对应于某个数 $x \geqslant 0$.

事实上, 根据记号 $\alpha_p, \cdots, \alpha_{p-n}, \cdots$ 可做出形如 (4) 式的数列 $\{r_n\}$. 由于 $r_0 \leqslant r_1 \leqslant \cdots \leqslant r_n \leqslant \cdots$, 再考虑到 (6) 式和 (7) 式, 就得到

$$r_0 \leqslant r_1 \leqslant \cdots \leqslant \cdots < \cdots \leqslant r_n + \frac{1}{q^{n-p}} \leqslant \cdots \leqslant r_1 + \frac{1}{q^{1-p}} \leqslant r_0 + \frac{1}{q^{-p}}. \tag{8}$$

以上关系中严格不等式的符号, 应这样来理解: 序列左边的任何元小于序列右边的任何元. 这是由 (7) 式推知的.

如果取

$$x = \sup_{n \in \mathbb{N}} r_n \left(= \inf_{n \in \mathbb{N}} \left(r_n + \frac{1}{q^{n-p}} \right) \right),$$

那么序列 r_n 将满足条件 (4),(5), 即记号 $\alpha_p, \cdots, \alpha_{p-n}, \cdots$ 对应于所得的数 $x \in \mathbb{R}$.

因此, 我们已把正数 x 一一地对应于这样的记号: 当 $p \geqslant 0$ 时, 对应到 $\alpha_p \cdots \alpha_0, \cdots$; 当 $p < 0$ 时, 对应到 $\underbrace{0.0 \cdots 0}_{|p|个零}\alpha_p \cdots$, 它叫做 x 的 *q进位记法*. 记号中所用到的数字叫*数码*, 数码相对于小数点的位置叫数码的*位*.

我们约定, 数 $x < 0$ 对应于正数 $-x$ 的记号联上一个减号, 而把数 0 对应于 $0.0 \cdots 0 \cdots$.

这样就建成了*实数的 q 进位记数系统*.

在日常生活中最常用的是十进位记数系统, 但是按照技术原理来说, 常用的则是二进位记数法 (在电子计算机中). 三进位与八进位记数法在计算技术部分中也有用, 但是用得不多.

公式 (4)、(5) 说明, 如果数 x 在 q 进位记法中只保留有限位数字 (或者, 如果愿意的话, 把其余的写成零), 那么, 这时 x 的近似值 (4) 的绝对误差, 不超过最后保留的那一位的一个单位.

根据这些观察, 我们能用第 b 小段所得公式, 去估计在算术运算中由于把精确值代之以形如 (4) 式的相应近似值所产生的误差.

这个注记也有一定的理论价值. 这就是, 如果按第 b 小段我们把实数 x 与它的 q 进位记法等同起来. 那么, 一旦学会了对 q 进位记号直接进行算术运算, 也就建立了一个新的实数模型. 看来, 从计算观点看, 这是最有价值的实数模型.

这样做时, 需要解决的基本问题如下.

需把两个 q 进位记号对应于一新的 q 进位记号 —— 原来两个记号的和. 当然, 它是一步一步构造出来的, 即把已知数的越来越精确的有理近似值相加, 随之得到它们的和的相应的有理近似值. 利用上面的注记, 可以证明随着两个加数的近似值精确程度的增加, 和数将有越来越多的 q 进位数码不再随着加数近似值的精确度的进一步增加而变化.

对于乘法也需要解决这样的问题.

另一种从有理数得出所有实数的不太具有构造性的方法是属于戴德金的.

戴德金把实数与有理数集 $\mathbb{Q}$ 的一个分割等同起来, 这就是把 $\mathbb{Q}$ 分成没有共同元素的两个集合 A, B, 使得 $\forall a \in A \forall b \in B (a < b)$. 这样处理实数时, 我们所采用的完备公理变成了著名的戴德金定理. 由于这个缘故, 常把我们所采用的完备公理称为戴德金公理.

这样, 在本节中引出了一些重要的数类. 指出了自然数与有理数的基础作用. 指出了怎样由我们采用的公理系统, 推出这些数的基本性质. 给出有关实数集的各种模型的基本知识. 讨论了实数理论的计算方面的问题: 当用近似值做算术运算时的误差估计; q 进位计数系统.

练　习

1. 依据归纳原理证明:

a) 实数之和 $x_1 + \cdots + x_n$ 与加括号时的位置无关;

b) 对乘积 $x_1 \cdot \cdots \cdot x_n$ 也如此;

c) $|x_1 + \cdots + x_n| \leqslant |x_1| + \cdots + |x_n|$;

d) $|x_1 \cdots x_n| = |x_1| \cdots |x_n|$;

e) $(m, n \in \mathbb{N}) \wedge (m < n) \Rightarrow ((n - m) \in \mathbb{N})$;

f) 当 $x > -1$ 且 $n \in \mathbb{N}$ 时, $(1 + x)^n \geqslant 1 + nx$; 同时, 只有 $n = 1$ 或 $x = 0$ 时等号成立 (伯努利不等式);

g) $(a + b)^n = a^n + \dfrac{n}{1!}a^{n-1}b + \dfrac{n(n-1)}{2!}a^{n-2}b^2 + \cdots + \dfrac{n(n-1)\cdots 2}{(n-1)!}ab^{n-1} + b^n$(牛顿二项式).

2. a) 验证 $\mathbb{Z}$ 与 $\mathbb{Q}$ 都是归纳集;

b) 举出一个异于 N、Z、Q、R 的归纳集的例子.

3. 试证, 任何归纳集无上界.

4. a) 任何归纳集是无穷集 (即它与自己的一个真子集等势).

b) 集 $E_n = \{x \in \mathbb{N} | x < n\}$ 是有限集 (用 n 来记 $\mathrm{card} E_n$).

5. a) 欧几里得辗转除法. 设 $m, n \in \mathbb{N}$ 且 $m > n$. 最大公约数($\mathrm{HCF}(m,n) = d \in \mathbb{N}$) 可用以下由一串带余除法构成的欧几里得辗转除法, 经有限步算出:

$$
\begin{aligned}
m &= q_1 n + r_1 \quad (r_1 < n),\\
n &= q_2 r_1 + r_2 \quad (r_2 < r_1),\\
r_1 &= q_3 r_2 + r_3 \quad (r_3 < r_2),\\
&\cdots\cdots\cdots\cdots\cdots\\
r_{k-1} &= q_{k+1} r_k + 0,
\end{aligned}
$$

而 $d = r_k$.

b) 如果 $d = \mathrm{HCF}(m,n)$, 那么, 必能找到 $p, q \in \mathbb{Z}$ 使得 $pm + qn = d$; 特别当 m, n 互素时, 则得 $pm + qn = 1$.

6. 试独立地证明算术基本定理 (公式在 §2 第 2.a 段)

7. 如果自然数的乘积 $m \cdot n$ 能被素数 p 整除, 即 $m \cdot n = p \cdot k, k \in \mathbb{N}$, 那么, 或 m 或 n 能被 p 整除.

8. 由算术基本定理推出素数集是无穷集.

9. 试证, 如果自然数 n 不是 k^m 这种形式, 这里的 $k, m \in \mathbb{N}$, 那么方程 $x^m = n$ 没有有理根.

10. 试证, 在任何 q 进位记数法中, 有理数必是循环的, 即从某一位开始, 由周期重复的一组数码构成.

11. 设 $\alpha \in \mathbb{R}$ 为无理数, 如果对于任何自然数 $n, N \in \mathbb{N}$, 存在有理数 $\frac{p}{q}$, 使 $\left|\alpha - \frac{p}{q}\right| < \frac{1}{Nq^n}$, 我们就说 α 能用有理数逼近得好.

a) 造一个用有理数逼近得好的无理数例子

b) 证明用有理数逼近得好的无理数, 不可能是代数数, 即它是超越数 (这是刘维尔(Liouville)① 定理).

12. 据分数定义知道 $\frac{m}{n} := m \cdot n^{-1}$, 这里 $m \in \mathbb{Z}, n \in \mathbb{N}$. 试导出分数加法, 乘法, 除法的"规则", 以及两个分数相等的条件.

13. 验证有理数集 $\mathbb{Q}$ 满足实数集的一切公理, 但完备公理除外.

14. 采用实数集的几何模型, 即数轴, 试说明在此模型中怎样作出 $a + b, a - b, ab, \frac{a}{b}$.

15. a) 在数轴上解释完备公理.

b) 证明上确界原理与完备公理等价.

① 刘维尔(Liouville)(1809—1882)—— 法国数学家, 从事复分析、几何、微分方程、数论、力学研究.

16. a) 如果 $A \subset B \subset \mathbb{R}$, 那么 $\sup A \leqslant \sup B$, 而 $\inf A \geqslant \inf B$.

b) 设 $\mathbb{R} \supset X \neq \varnothing$, 且 $\mathbb{R} \supset Y \neq \varnothing$. 若 $\forall x \in X, \forall y \in Y$ 满足 $x \leqslant y$. 那么 X 上有界而 Y 下有界, 并且 $\sup X \leqslant \inf Y$.

c) 如果 b) 中的 X 与 Y 又满足 $X \cup Y = \mathbb{R}$, 那么, $\sup X = \inf Y$.

d) 如果 X, Y 是 c) 中所定义的集, 那么, 或者 $\exists \max X$, 或者 $\exists \min Y$(这是戴德金定理).

e) (接 d)) 试证戴德金定理与完备公理等价.

17. 设 $A+B$ 是形如 $a+b$ 的数的集合, $A \cdot B$ 是形如 $a \cdot b$ 的数的集合, 其中 $a \in A \subset \mathbb{R}, b \in B \subset \mathbb{R}$. 试检查是否总有

a) $\sup(A+B) = \sup A + \sup B$;

b) $\sup(A \cdot B) = \sup A \cdot \sup B$.

18. 设 $-A$ 是形如 $-a$ 的数的集合, 这里 $a \in A \subset \mathbb{R}$, 试证, $\sup(-A) = -\inf A$.

19. a) 试证方程 $x^n = a$ 当 $n \in \mathbb{N}$ 且 $a > 0$ 时有正根 (记作 $\sqrt[n]{a}$ 或 $a^{\frac{1}{n}}$),叫n次算术根.

b) 验证, 当 $a > 0, b > 0$ 且 $n, m \in \mathbb{N}$ 时,

$$\sqrt[n]{ab} = \sqrt[n]{a} \cdot \sqrt[n]{b} \quad \text{且} \quad \sqrt[n]{\sqrt[m]{a}} = \sqrt[mn]{a}.$$

c) $(a^{\frac{1}{n}})^m = (a^m)^{\frac{1}{n}} =: a^{\frac{m}{n}}$ 而 $a^{\frac{1}{n}} \cdot a^{\frac{1}{m}} = a^{\frac{1}{n}+\frac{1}{m}}$.

d) $(a^{\frac{m}{n}})^{-1} = (a^{-1})^{\frac{m}{n}} =: a^{-\frac{m}{n}}$.

e) 试证, 对于任何 $r_1, r_2 \in \mathbb{Q}$

$$a^{r_1} \cdot a^{r_2} = a^{r_1+r_2} \quad \text{且} \quad (a^{r_1})^{r_2} = a^{r_1 r_2}.$$

20. a) 试证集合间的包含关系是偏序关系 (但不是全序关系!).

b) 设 A, B, C 满足 $A \subset C, B \subset C, A \backslash B \neq \varnothing$ 且 $B \backslash A \neq \varnothing$. 像 a) 中那样, 在这三个集间引进偏序. 试指出集合 $\{A, B, C\}$ 中的极大元与极小元 (注意其不唯一性!).

21. a) 试证, 像有理数集 $\mathbb{Q}$ 一样, 形如 $a + b\sqrt{n}$ 之集 $\mathbb{Q}(\sqrt{n})$ 是有序域, 其中 $a, b \in \mathbb{Q}$, 而 n 是不等于平方整数的固定的自然数. $\mathbb{Q}(\sqrt{n})$ 还满足阿基米德原理, 但不满足完备公理.

b) 如果在 $\mathbb{Q}(\sqrt{n})$ 中保留以前的算术运算, 而按

$$(a + b\sqrt{n} \leqslant a' + b'\sqrt{n}) := ((b \leqslant b') \vee ((b = b') \wedge (a \leqslant a')))$$

规定序关系, 检验实数公理中的哪一些对 $\mathbb{Q}(\sqrt{n})$ 不再满足. 这时对于 $\mathbb{Q}(\sqrt{n})$ 来说, 阿基米德原理是否还对?

c) 在有理系数或实系数多项式集合 $\mathbb{P}[x]$ 中建立序关系使

$$a_0 + a_1 x + \cdots + a_m x^m \succ 0, \text{若} a_m > 0.$$

d) 试证, 系数 a_i, b_j 属于 $\mathbb{Q}$ 或属于 $\mathbb{R}$ 的所有有理分式

$$\mathrm{R}_{m,n} = \frac{a_0 + a_1 x + \cdots + a_m x^m}{b_0 + b_1 x + \cdots + b_n x^n}$$

的集合 $\mathbb{Q}[x]$, 在按 $\mathrm{R}_{m,n} \succ 0$, 若 $\dfrac{a_m}{b_n} > 0$ 引入序关系与通常的算术运算之后, 构成有序域, 但不是阿基米德序域. 这意味着, 阿基米德原理不能抛开完备公理从 $\mathbb{R}$ 的其他公理推出.

22. 设 $n \in \mathbb{N}$ 且 $n > 1$. 在集 $E_n = \{0, 1, \cdots, n-1\}$ 中定义二数之和与积为将该二数在 $\mathbb{R}$ 中的"普通"的和与积除以 n 所得的余数. 在 E_n 中如此定义了加乘运算后记作 $\mathbb{Z}_n$.

a) 试证, 如果 n 不是素数, 那么在 $\mathbb{Z}_n$ 中存在着异于零的数 m, k, 使得 $m \cdot k = 0$ (这样的数叫做零因子) 这就是说, 由 $a \cdot b = c \cdot b$, 即使 $b \neq 0$, 在 $\mathbb{Z}_n$ 中也推不出 $a = c$ 来.

b) 试证, 当 p 为素数时, $\mathbb{Z}_p$ 中没有零因子, 而 $\mathbb{Z}_p$ 是域.

c) 试证, 对任何素数 p, 域 $\mathbb{Z}_p$ 不能给以一种顺序, 使它与 $\mathbb{Z}_p$ 中的算术运算相协调.

23. 设 $\mathbb{R}$ 与 $\mathbb{R}'$ 是实数集的两个模型, $f : \mathbb{R} \to \mathbb{R}'$ 是对于任何 $x, y \in \mathbb{R}$ 满足

$$f(x+y) = f(x) + f(y), f(x \cdot y) = f(x) \cdot f(y)$$

的映射. 试证:

a) $f(0) = 0'$;

b) 如果 $f(x) \not\equiv 0$, 则 $f(1) = 1'$(我们以后认为这个条件成立);

c) $f(m) = m'$, 这里 $m \in \mathbb{Z}$ 而 $m' \in \mathbb{Z}'$, 并且 $f : \mathbb{Z} \to \mathbb{Z}'$ 是双射且保序;

d) $f\left(\dfrac{m}{n}\right) = \dfrac{m'}{n'}$, 这里 $m, n \in \mathbb{Z}, n \neq 0, m', n' \in \mathbb{Z}', n' \neq 0', f(m) = m', f(n) = n'$. 因此 $f : \mathbb{Q} \to \mathbb{Q}'$ 是保序双射;

e) $f : \mathbb{R} \to \mathbb{R}'$ 是保序双射.

24. 据上题及完备公理, 试证实数集, 在同构 (实现方法) 意义下, 由实数集的公理系统完全确定, 即如果 $\mathbb{R}$ 与 $\mathbb{R}'$ 是满足 (实数) 公理的两个集合, 那么, 存在双方单值映射 $f : \mathbb{R} \to \mathbb{R}'$, 且保持算术运算和序关系:

$$f(x+y) = f(x) + f(y), f(x \cdot y) = f(x) \cdot f(y) \text{以及} (x \leqslant y) \Leftrightarrow (f(x) \leqslant f(y)).$$

25. 在电子计算机中, 数 x 表成

$$x = \pm q^p \sum_{n=1}^{k} \frac{\alpha_n}{q^n}$$

的形式, 其中 p 是 x 的阶数, $M = \displaystyle\sum_{n=1}^{k} \frac{\alpha_n}{q^n}$ 是 x 的尾数$\left(\dfrac{1}{q} \leqslant M < 1\right)$.

同时, 机器只能对一定范围内的数进行运算: 当 $q = 2$ 时, 通常 $|p| \leqslant 64$, 而 $k = 35$. 试估计在十进位制中机器的工作范围.

26. a) 对六进位制列出一个 (范围为 6×6) 乘法表来.

b) 利用问题 a) 的结果, 做六进位制的"竖式"乘法:

$$\begin{array}{r} (532)_6 \\ \times(145)_6 \\ \hline \end{array}$$

再用十进位制计算, 以验证你算的结果.

c) 以"角"式做除法:

$$(1\,301)_6 \,|\, \underline{(25)_6}$$

再用十进位制计算, 以验证之.

d) 做"竖"式加法

$$\begin{array}{r} (4\,052)_6 \\ +(3\,125)_6 \\ \hline \end{array}$$

27. 将 $(100)_{10}$ 写成二进位及三进位数的形式.

28. a) 试证: 与整数能唯一地写成

$$(\alpha_n\alpha_{n-1}\cdots\alpha_0)_3 \quad \alpha_i \in \{0,1,2\}$$

之形式的同时, 它还能唯一地写成

$$(\beta_n\beta_{n-1}\cdots\beta_0)_3 \quad \beta_j \in \{-1,0,1\}.$$

的形式.

b) 一堆硬币中有一枚是假的, 它只是在重量上与其他的不同. 若限于在天秤上最多称三次就能挑出其中的假币来, 问题这堆硬币最多是多少枚?

29. 为了探明七位数电话号中的任一个, 最少需要给出多少次"是"与"否"的回答?

30. a) 利用二十个十进数码 (例如, 全部十个数码每一个都占两位). 可以给出多少个不同的数? 对于二进位制回答同一问题, 从比较结果看, 这两种进位制哪一种比较经济?

b) 试估计, 由 q 进制的 n 个号码能写出多少个不同的数来.(答案:$(q)^{\frac{n}{q}}$.)

c) 设自变量 x 为自然数, 作出函数 $f(x)=x^{\frac{n}{x}}$ 的图像, 并比较各种计数法的经济性.

§3. 与实数集的完备性有关的基本引理

在这里, 我们建立几个简单、有益的原理, 其中每一个都能当作完备公理而成为建立实数理论的基础①.

我们之所以称这些原理为基本引理是因为, 它们在分析学定理的各种各样的证明中有广泛应用.

1. 闭区间套引理 (柯西–康托尔原理)

定义 1 以自然数为自变量的函数 $f:\mathbb{N}\to X$ 叫做*序列*, 更完整的说法是*集合 X 中的元素序列*.

与 $n\in\mathbb{N}$ 相对应的、函数 f 的值 $f(n)$, 常记作 x_n, 并称为序列的第 n 项.

定义 2 设 $X_1, X_2,\cdots,X_n,\cdots$ 是集合的序列. 如果

$$X_1\supset X_2\supset\cdots\supset X_n\supset\cdots, 即\forall n\in\mathbb{N}(X_n\supset X_{n+1}),$$

那么, 就说它是*集列套*.

① 参看节末问题 4.

引理 *对于任何闭区间套**

$$I_1 \supset I_2 \supset \cdots \supset I_n \supset \cdots,$$

存在一点 $c \in \mathbb{R}$, 属于这些闭区间的每一个.

此外, 如果对于任何 $\varepsilon > 0$, 在序列中能找到闭区间 I_k, 使其长 $|I_k| < \varepsilon$, 那么 c 就是所有闭区间的唯一公共点.

◀ 首先注意, 对我们序列中的任何两个 $I_m = [a_m, b_m], I_n = [a_n, b_n]$, 必有 $a_m \leqslant b_n$. 因为, 如若不然, 就得到 $a_n \leqslant b_n < a_m \leqslant b_m$, 即闭区间 I_m, I_n 没有公共点, 但同时其中下标较大的一个又要包含在另一个之中.

因此数集 $A = \{a_m, m \in \mathbb{N}\}, B = \{b_n, n \in \mathbb{N}\}$ 满足完备公理的条件. 据此存在一数 $c \in \mathbb{R}$ 使得 $\forall a_m \in A, \forall b_m \in B$ 满足 $a_m \leqslant c \leqslant b_n$. 尤其是对于任何 $n \in \mathbb{N}$ 有 $a_n \leqslant c \leqslant b_n$. 这就是说, 点 c 位于一切闭区间 I_n 中.

今设 c_1 与 c_2 为具有这种性质的两个点. 如果它们不同, 比如说 $c_1 < c_2$, 那么, 对于任何 $n \in \mathbb{N}, a_n \leqslant c_1 < c_2 \leqslant b_n$. 所以 $0 < c_2 - c_1 < b_n - a_n$, 因而在我们的闭区间序列中, 每个闭区间的长度不能小于 $c_2 - c_1$. 这就是说, 如果在序列中有长度任意短的闭区间, 那么它们的公共点必定是唯一的. ▶

2. 有限覆盖引理 (博雷尔–勒贝格[①]原理)

定义 3 设 $S = \{X\}$ 是由一些集合 X 构成的集族. 如果 $Y \subset \bigcup\limits_{X \in S} X$(即集合 Y 的每个元素 y, 至少含在集族 S 中的一个集合 X 中), 就说 *S覆盖集合 Y*.

集合 $S = \{X\}$ 的子集也是一个集族, 叫做 S 的*子族*. 于是, 集族的子族本身还是同一类型的集族.

引理 *在覆盖一个闭区间的任何开区间族中, 存在着覆盖这一闭区间的有限子族.*

◀ 设 $S = \{U\}$ 是覆盖闭区间 $[a, b] = I_1$ 的开区间族, U 是开区间. 假如闭区间 I_1 不能用 S 中的有限个来覆盖, 那么, 把 I_1 等分成两个闭区间时, 至少有一个不能有有限覆盖, 将这个闭区间记作 I_2. 对 I_2 进行同样的对半分手续就得到 I_3, 等等.

这样, 就得到一个闭区间套

$$I_1 \supset I_2 \supset \cdots \supset I_n \supset \cdots,$$

它们都不能用开区间族 S 的有限子族覆盖. 因为这样做 n 步时所得到的线段 I_n 之长为 $|I_n| = \dfrac{|I_1|}{2^n}$, 所以, 在序列 $\{I_n\}$ 中有长度任意小的闭区间 (参看 §2,4c 之引理).

① 博雷尔(Borel) (1871—1956), 勒贝格(A.Lebesgue) (1875—1941)—— 著名的法国数学家, 函数论专家.

* 译者注. 请读者注意, 根据定义 (见上一节定义 6) 闭区间, 与无界区间不同, 是有界集

据闭区间套引理, 存在一点 c, 属于每一个闭区间 $I_n, n\in\mathbb{N}$. 因为 $c\in I_1=[a,b]$, 所以必有族 S 中的一个开区间 $]\alpha,\beta[=U\in S$ 包含着点 c, 即 $\alpha<c<\beta$. 令 $\varepsilon=\min\{c-\alpha,\beta-c\}$, 在所做出的那一列闭区间中找出一个 I_n, 使 $|I_n|<\varepsilon$. 因为 $c\in I_n$, 且 $|I_n|<\varepsilon$, 因此 $I_n\subset U=]\alpha,\beta[$. 但这与 I_n 不能由开区间族中的有限个所覆盖这件事矛盾. ►

3. 极限点引理 (波尔察诺–魏尔斯特拉斯[①]原理) 我们记得, 含有点 $x\in\mathbb{R}$ 的开区间, 我们称它为 x 的邻域; 而开区间 $]x-\delta,x+\delta[$ 叫做点 x 的 δ 邻域.

定义 4 假如点 $p\in\mathbb{R}$ 的任何邻域都包含 $X\subset\mathbb{R}$ 的一个无穷子集, 就称点 $p\in\mathbb{R}$ 为集合 X 的*极限点*.

这个条件显然与以下的条件等价: 在点 p 的任何邻域中, 至少含有 X 中的一个不与 p 重合的点 (请验证!).

举几个例子

设 $X=\left\{\dfrac{1}{n}\in\mathbb{R}\,|\,n\in\mathbb{N}\right\}$, 则只有点 $0\in\mathbb{R}$ 是 X 的极限点.

对于开区间 $]a,b[$ 来说, 闭区间 $[a,b]$ 的每一点是极限点, 并且不再有其他的极限点.

$\mathbb{R}$ 的每一点是有理数集 $\mathbb{Q}$ 的极限点, 因为我们知道, 在实数集的任何开区间中都含有有理数.

引理 (波尔察诺–魏尔斯特拉斯) *每个无穷有界集至少有一个极限点.*

◄ 设 X 是给定的 $\mathbb{R}$ 的子集. 由 X 有界性的定义推知, X 含在某个闭区间 $[a,b]=I\subset\mathbb{R}$ 中, 现在证明, 闭区间 I 中至少有 X 的一个极限点.

如果不如此, 那么每点 $x\in I$ 有邻域 $U(x)$, 其中或者不含 X 的点, 或者只有有限多个这种点. 对于每个 $x\in I$ 所构造的这些邻域的总体 $\{U(x)\}$ 形成闭区间 I 的一个开区间覆盖; 根据有限覆盖引理, 能从其中取出有限个开区间 $U(x_1),\cdots,U(x_n)$ 来, 仍然覆盖闭区间 I. 但是 $X\subset I$, 所以这有限个区间也覆盖着 X. 然而在每个 $U(x_i)$ 中只含 X 的有限个点, 所以在它们的并中也只含 X 的有限个点, 即 X 是有限集. 这与假设矛盾, 从而完成了定理的证明. ►

练　习

1. 试证

a) 如果 I 是任意一个闭区间套, 那么

$$\sup\{a\in\mathbb{R}\,|\,[a,b]\in I\}=\alpha\leqslant\beta=\inf\{b\in\mathbb{R}\,|\,[a,b]\in I\},$$

① 波尔察诺 (B. Bolzano) (1781—1848) —— 捷克数学家与哲学家. 魏尔斯特拉斯 (K. Weierstrass) (1815—1897) —— 德国数学家, 他非常关注数学分析的逻辑基础.

且

$$[\alpha, \beta] = \bigcap_{[a,b]\in I} [a, b].$$

b) 如果 I 是一个开区间套, 那么, 交集 $\bigcap\limits_{]a,b[\in I}]a, b[$ 可能是空集.

提示: $]a_n, b_n[=]0, \frac{1}{n}[$.

2. 试证

a) 覆盖一个闭区间的闭区间族不必包含此闭区间的有限子覆盖.

b) 覆盖一个开区间的开区间族不必包含此开区间的有限子覆盖.

c) 覆盖开区间的闭区间族也不必包含此开区间的有限子覆盖.

3. 试证, 如果将所有实数之集 $\mathbb{R}$ 代之以有理数之集 $\mathbb{Q}$, 把闭区间, 开区间与点 $r \in \mathbb{Q}$ 之邻域理解成 $\mathbb{Q}$ 的相应的子集, 那么, 上面所证明的三个基本引理的任一个都不再成立.

4. 试证, 如果把

a) 波尔察诺–魏尔斯特拉斯原理或

b) 博雷尔–勒贝格原理

取作为实数集的完备公理, 那么就得到与前面等价的、$\mathbb{R}$ 的公理系统.

提示: 由 a) 推出阿基米德原理及原来形式的完备公理.

c) 用柯西–康托尔原理代替公理系统中的完备公理, 另外, 再假定阿基米德原理, 这样得到的公理系统与原公理系统等价 (参看上节问题 21).

§4. 可数集与不可数集

现在, 我们对于在第一章中已经谈到的有关集合的知识, 做些不大的但对以后很有用的补充.

1. 可数集

定义 1 如果集合 X 与自然数集 $\mathbb{N}$ 等势, 即 $\operatorname{card} X = \operatorname{card} \mathbb{N}$, 就称 X 为可数集.

命题 a) *可数集的无穷子集是可数集.*

b) *有限个或可数个可数集的并是可数集.*

◀ a) 只需验证自然数集 $\mathbb{N}$ 的每个无穷子集 E 与 $\mathbb{N}$ 等势. 我们用下面的方式来建立所需要的双射 $f : \mathbb{N} \to E$. E 有最小元, 使它与 $1 \in \mathbb{N}$ 对应, 并记作 $e_1 \in E$. 集 E 是无穷集, 所以 $E_2 = E \setminus \{e_1\}$ 不空, 使 E_2 的最小元与数 2 对应, 记作 $e_2 \in E_2$. 然后考察 $E_3 = E \setminus \{e_1, e_2\}$, 等等. 因为 E 是无穷集, 所以这种做法不能到第 n 步就完了. 因此, 据归纳原理, 对于每个 $n \in \mathbb{N}$, 有某个数 $e_n \in E$ 与之对应, 所建立的映射 $f : \mathbb{N} \to E$ 显然是内射.

现在只需验证映射 f 是满射, 即 $f(\mathbb{N})=E$. 设 $e\in E$. 集合 $\{n\in\mathbb{N}|n\leqslant e\}$ 是有限集, 它的子集 $\{n\in E|n\leqslant e\}$ 更是有限集了. 设在此子集里有 k 个元素. 这时, 根据我们的做法, $e=e_k$.

b) 如果 $X_1,\cdots,X_n,\cdots$ 是可数集族, 并且每个集合

$$X_m=\{x_m^1,\cdots,x_m^n,\cdots\}$$

本身也是可数集, 那么, 由元素 $x_m^n(m,n\in\mathbb{N})$ 组成的集 $X=\bigcup\limits_{m\in\mathbb{N}}X_m$ 的势不小于每个 X_m 的势, 所以 X 是无穷集. 可以把元素 $x_m^n\in X_m$ 与自然数序对 (m,n) 等同起来. 因此, X 的势不大于这些序对之集的势. 然而用公式

$$(m,n)\mapsto\frac{(m+n)(m+n+1)}{2}+m$$

所给出的映射 $f:\mathbb{N}\times\mathbb{N}\to\mathbb{N}$, 很容易验证为双射. (它的直观意义是: 我们这样来给平面上坐标为 (m,n) 的点编号, 从 $m+n$ 为常数的一条对角线, 过渡到另一条对角线, 在这后一条对角线上, $m+n$ 比前一条对角线多 1.)

因此, 自然数序对是可数的. 因而 $\operatorname{card}X\leqslant\operatorname{card}\mathbb{N}$. 又因 X 是无穷集, 根据 a), 推知 $\operatorname{card}X=\operatorname{card}\mathbb{N}$. ►

由以上命题推知, 可数集的任何子集, 或有限, 或可数. 如果知道了某个集合或为有限, 或为可数, 就说它是*至多可数集* (等价的写法:$\operatorname{card}X\leqslant\operatorname{card}\mathbb{N}$).

特别地, 现在可以断定,*至多可数个至多可数集的并是至多可数的*.

推论 1) $\operatorname{card}\mathbb{Z}=\operatorname{card}\mathbb{N}$.

2) $\operatorname{card}\mathbb{N}^2=\operatorname{card}\mathbb{N}$.

这个结论说的是, 可数集的直积是可数集

3) $\operatorname{card}\mathbb{Q}=\operatorname{card}\mathbb{N}$, *即有理数集是可数的*.

◄ 有理数 $\dfrac{m}{n}$ 由整数序对 (m,n) 给出.

两个序对 $(m,n),(m',n')$ 给出同一有理数的充要条件是它们成比例. 因此, 对每个有理数 r, 可取 $n\in\mathbb{N}$ 使存在 $m\in\mathbb{Z},\dfrac{m}{n}=r$ 且 n 是满足这些条件的最小者. 易见, 对每个 $r\in\mathbb{Q}$, 这样的整数对 (m,n) 是唯一的. 于是, 得知, 集 $\mathbb{Q}$ 等势于集 $\mathbb{Z}\times\mathbb{Z}$ 的某个无穷子集. 但是 $\operatorname{card}\mathbb{Z}^2=\operatorname{card}\mathbb{N}$, 因此, $\operatorname{card}\mathbb{Q}=\operatorname{card}\mathbb{N}$. ►

4) *代数数集是可数集*

◄ 首先注意, 由 $\operatorname{card}\mathbb{Q}\times\mathbb{Q}=\operatorname{card}\mathbb{N}$, 据归纳法得知, 对任何 $k\in\mathbb{N}$, 有 $\operatorname{card}\mathbb{Q}^k=\operatorname{card}\mathbb{N}$.

元素 $r\in\mathbb{Q}^k$ 是 k 个有理数的序组 $(r_1,\cdots,r_k)$.

有理系数的 k 次代数方程可以写成

$$x^k+r_1x^{k-1}+\cdots+r_k=0$$

的形式, 其中最高次幂的系数等于 1. 这样, 有多少个不同的 k 次代数方程, 就有多少个有理数序组, 即 k 次有理系数代数方程构成可数集.

一切有理系数代数方程也组成可数集, 因为它是可数集的可数并 (k 次有理系数代数方程之集关于次数 $k \in \mathbb{N}$ 的并). 每个这样的方程只有有限个根, 因此, 代数数之集最多可数, 然而它是无限集, 所以必是可数集. ►

2. 连续统的势

定义 2 实数集 $\mathbb{R}$ 也叫做数的*连续统*[①], 而它的势叫做*连续统的势*.

定理 (康托尔) $\operatorname{card} \mathbb{N} < \operatorname{card} \mathbb{R}$.

定理断定无穷集 $\mathbb{R}$ 的势大于无穷集 $\mathbb{N}$ 的势.

◄ 我们来证闭区间 $[0,1]$ 上的点就已是不可数的了.

假如它可数, 即它能写成序列 $x_1, x_2, \cdots, x_n, \cdots$ 的形式. 在闭区间 $[0,1] = I_0$ 上, 取定一闭区间 I_1, 使 I_1 之长不为零, 且 I_1 不含 x_1; 在 I_1 中取定闭区间 I_2, 使 $|I_2| \neq 0$, 并且 I_2 不含点 x_2; 如果已经做出了闭区间 I_n 使 $|I_n| \neq 0$, 且 I_n 不包含 x_n, 那么, 从其中即可截取 I_{n+1} 使 $|I_{n+1}| \neq 0$, 且 $x_{n+1} \notin I_{n+1}$. 据闭区间套原理, 存在一点 c, 它属于一切闭区间 $I_0, I_1, \cdots, I_n, \cdots$. 但是按照我们的做法, 闭区间 $I_0 = [0,1]$ 的这个点, 不能是点列 $x_1, x_2, \cdots, x_n, \cdots$ 中的任何一点. ►

推论 1) *$\mathbb{Q} \neq \mathbb{R}$ 且无理数存在.*

2) *因为代数数之集可数, 所以存在超越数.*

(在解答了本节末的问题 3 以后, 读者大概希望把上面的断语改成:"在实数集中, 代数数也是有时会遇到的数*".)

早在集合论一开始, 就出现了是否存在着可数集势与连续统势的中间势的集这个问题, 并提出了叫做 "连续统假定" 的命题, 它断定不存在中间势.

这个问题深刻地涉及数学的基础, 它被美国数学家科恩 (P. Cohen) 在 1963 年完全解决. 科恩证明了连续统假定是不可解的, 并指出, 无论是它本身还是它的否定与公理化集合论都没有矛盾, 所以连续统假定在这公理系统框架内既不能被证明, 也不能被否定. 这与欧几里得关于平行线的第五公设独立于其他几何公理的情况完全类似.

练 习

1. 试证所有实数之集与开区间 $]-1,1[$ 之点的集等势.
2. 在以下各对点集之间直接建立双方单值对应关系.

---

① Continuum (拉丁文)—— 连续的, 不间断的.

* 译者注: 意指在实数集中并不经常遇到代数数.

a) 两个开区间;

b) 两个闭区间;

c) 闭区间与开区间;

d) 闭区间 $[0,1]$ 与实数集 $\mathbb{R}$.

3. 试证

a) 任何无穷集含有可数子集;

b) 偶数集与所有自然数之集 $\mathbb{N}$ 等势;

c) 无穷集与一个至多可数集的并的势与原来那个无穷集的势一样;

d) 无理数集具有连续统势;

e) 超越数之集具有连续统势.

4. 试证

a) 递增自然数列之集与形如 $0\cdot\alpha_1\alpha_2\cdots$ 之二进小数之集等势;

b) 可数集的一切子集之集合, 具有连续统的势.

5. 试证

a) 集合 X 的一切子集的集合 $\mathcal{P}(X)$, 与一切在 X 上定义而取值为 0 或 1 的函数 (即一切映射 $f: X \to \{0,1\}$) 之集等势;

b) 设 X 是 n 元素之有限集, 那么 $\text{card}\ \mathcal{P}(X) = 2^n$;

c) 注意到问题 4b) 与 5a) 之结论, 可以得到 $\text{card}\ \mathcal{P}(X) = 2^{\text{card}\ X}$, 特别地, $\text{card}\ \mathcal{P}(\mathbb{N}) = 2^{\text{card}\ \mathbb{N}} = \text{card}\ \mathbb{R}$;

d) 对于任何集合 X, 有

$$\text{card}\ X < 2^{\text{card}X},$$

特别地, 对一切 $n \in \mathbb{N}$, 有 $n < 2^n$.

提示: 参看第一章 §4 第 1 段的康托尔定理.

6. 设 $X_1, \cdots, X_m$ 是有限个有限集, 试证

$$\begin{aligned}\text{card}\left(\bigcup_{i=1}^{m} X_i\right) = &\sum_{i_1} \text{card}\ X_{i_1} - \sum_{i_1<i_2} \text{card}(X_{i_1} \cap X_{i_2}) \\ &+ \sum_{i_1<i_2<i_3} \text{card}(X_{i_1} \cap X_{i_2} \cap X_{i_3}) - \cdots \\ &+(-1)^{m-1}\text{card}(X_1 \cap \cdots \cap X_m),\end{aligned}$$

并且求和是对于在 $1, \cdots, m$ 范围内满足和号下的不等式的一切可能的选取来做的.

7. 在闭区间 $[0,1] \subset \mathbb{R}$ 上, 画出那样的数 x 的集合:$x \in [0,1]$, 其三进位记法 $x = 0.\alpha_1\alpha_2\alpha_3\cdots$ ($\alpha_i \in \{0,1,2\}$) 具下列三性质之一:

a) $\alpha_1 \neq 1$;

b) $(\alpha_1 \neq 1) \wedge (\alpha_2 \neq 1)$;

c) $\forall i \in \mathbb{N}(\alpha_i \neq 1)$(康托尔集).

8. (续问题 7) 试证

a) 在二进位记法中不含有 1 的那些数 $x \in [0,1]$ 的集, 与二进位记法中表示成 $0.\beta_1\beta_2,\cdots$ 的所有数的集等势;

b) 康托尔集与闭区间 $[0,1]$ 的所有点之集等势.

第三章　极限

在对实数概念的全面讨论中, 我们特别指出了, 在测量实际物理量时, 得到的是它们的近似值序列, 随后还必须对这些近似值序列进行研究.

这种情况马上会产生至少三个问题:

1) 得到的这个近似序列与被测的量之间有什么联系? 我们从数学上考虑这个问题, 就是说, 我们希望准确地知道,"近似值序列" 这个词一般是什么意思; 它能在何种程度上刻画被测量的值; 这种刻画是否唯一, 或者说, 同一序列能对应着所测量的不同值吗?

2) 近似值的运算与精确值的同样的运算有什么关系, 能把精确值换成近似值的运算的特征是什么?

3) 如何根据数列本身来确定它是不是能任意逼近某个量的数列?

函数极限概念 —— 这是分析的基本概念之一 —— 对这些问题以及一些与之相近的问题提供了答案.

我们从考察以自然数为自变量的函数 (序列) 开始来阐述极限理论, 因为已经弄清了这些函数的基本作用, 而且, 实际上, 极限论的所有基本事实, 在这个最简单的情形, 就已经能清楚地看到了.

§1. 序列的极限

1. 定义和例子　我们回想一下下面的

定义 1　定义域为自然数集的函数 $f: \mathbb{N} \to X$ 叫做序列.

函数 f 的值 $f(n)$ 叫做序列的项, 它们用映射到达集 X 的元素, 并附加自变

量作为下标来表示: $x_n := f(n)$. 因此, 序列本身用符号 $\{x_n\}$ 表示, 也可写成 $x_1, x_2, \cdots, x_n, \cdots$ 的形式并称之为 *X* 中的序列或集合 *X* 中元素列.

元素 x_n 叫做序列的第 n 项.

在下面的几节中, 我们只讨论实数列 $f: \mathbb{N} \to \mathbb{R}$.

定义 2 称数 $A \in \mathbb{R}$ 为数列 $\{x_n\}$的极限, 如果对于点 A 的任何邻域 $V(A)$, 存在号码 N (其选取与 $V(A)$ 有关), 使得数列之所有标号大于 N 的项, 包含在点 A 的这个邻域 $V(A)$ 之中.

下面将对这个定义引入形式逻辑的记法, 然而, 我们先指出数列极限定义的另一种流行的表述:

称数 $A \in \mathbb{R}$ 为数列 $\{x_n\}$的极限, 如果对于任何 $\varepsilon > 0$, 存在着号码 N, 使得对于一切 $n > N$, 有 $|x_n - A| < \varepsilon$.

如果注意到点 A 的任何邻域 $V(A)$ 都包含这点的一个 ε 邻域, 就容易验证这两种说法的等价性 (请验证).

极限定义的这后一种说法是说, 不论给出怎样的精确度 $\varepsilon > 0$, 总能找到号码 N, 使以数列 $\{x_n\}$ 的项去逼近数 A 时, 只要 $n > N$, 绝对误差就小于 ε.

现在用逻辑符号写出极限的形式化的定义, 并约定用式子 "$\lim\limits_{n\to\infty} x_n = A$" 表示数列 $\{x_n\}$ 的极限是 A. 于是

$$\boxed{\left(\lim_{n\to\infty} x_n = A\right) := \forall V(A) \exists N \in \mathbb{N} \forall n > N (x_n \in V(A)).}$$

而相应地有

$$\left(\lim_{n\to\infty} x_n = A\right) := \forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n > N (\mid x_n - A \mid < \varepsilon).$$

定义 3 如果 $\lim\limits_{n\to\infty} x_n = A$, 就说数列 $\{x_n\}$收敛于A 或说趋于 A, 并记成 "当 $n \to \infty$ 时 $x_n \to A$". 有极限的序列叫做收敛数列. 没有极限的序列叫做发散数列.

讨论几个例子.

例 1 $\lim\limits_{n\to\infty} \dfrac{1}{n} = 0$. 因为当 $n > N = \left[\dfrac{1}{\varepsilon}\right]$ [①] 时, $\left|\dfrac{1}{n} - 0\right| = \dfrac{1}{n} < \varepsilon$, 成立.

例 2 $\lim\limits_{n\to\infty} \dfrac{n+1}{n} = 1$. 因为当 $n > N = \left[\dfrac{1}{\varepsilon}\right]$ 时, $\left|\dfrac{n+1}{n} - 1\right| = \dfrac{1}{n} < \varepsilon$ 成立.

例 3 $\lim\limits_{n\to\infty} \left(1 + \dfrac{(-1)^n}{n}\right) = 1$. 因为当 $n > N = \left[\dfrac{1}{\varepsilon}\right]$ 时, $\left|\left(1 + \dfrac{(-1)^n}{n}\right) - 1\right| = \dfrac{1}{n} < \varepsilon$ 成立.

① $[x]$ 是 x 的整数部分 (参看第二章 §2 第 3 段阿基米德原理的推论 10°).

例 4 $\lim\limits_{n\to\infty}\dfrac{\sin n}{n}=0$. 因为当 $n>N=\left[\dfrac{1}{\varepsilon}\right]$ 时, $\left|\dfrac{\sin n}{n}-0\right|\leqslant\dfrac{1}{n}<\varepsilon$ 成立.

例 5 $\lim\limits_{n\to\infty}\dfrac{1}{q^n}=0$, 如果 $|q|>1$.

我们根据极限的定义来验证它. 像在第二章 §2 第 4c 小节中所证的那样, 对于任何 $\varepsilon>0$, 能够找到一个数 $N\in\mathbb{N}$ 使得 $\dfrac{1}{|q|^N}<\varepsilon$. 因为 $|q|>1$, 所以对于任何 $n>N$ 将有

$$\left|\frac{1}{q^n}-0\right|=\frac{1}{|q|^n}<\frac{1}{|q|^N}<\varepsilon,$$

于是极限定义被满足了.

例 6 数列 $1,2,\dfrac{1}{3},4,\dfrac{1}{5},6,\dfrac{1}{7},\cdots$, 其第 n 项 $x_n=n^{(-1)^n},n\in\mathbb{N}$ 是发散的.

实际上, 如果 A 是数列的极限, 根据极限的定义, 在 A 的任何邻域内, 至多有有限项不含于其中.

数 $A\neq0$, 不可能是给定序列的极限. 因为取 $\varepsilon=\dfrac{|A|}{2}>0$, 则在 A 的 ε 邻域之外, 有数列中一切形如 $\dfrac{1}{2k+1}$ 的项, 这里 k 是满足 $\dfrac{1}{2k+1}<\dfrac{|A|}{2}$ 的自然数.

数A=0也不可能是这个数列的极限, 因为, 在以0为中心的单位邻域之外, 显然也有序列的无穷多项.

例 7 类似地可验证数列 $1,-1,1,-1,\cdots(x_n=(-1)^n)$ 没有极限.

2. 数列极限的性质

a. 普遍性质 下面得到的一组性质, 虽然只是对数列讨论的, 但是我们将来可以看到, 这些性质, 不止对数列成立, 而且有普遍的意义.

一个数列, 如果它只取一个值, 就叫做*常数列*.

定义 4 如果存在一个数 A 与号码 N, 使得当 $n>N$ 时 $x_n=A$, 就说数列 $\{x_n\}$ 为*最终常数列*.

定义 5 如果存在一数 M, 使得对于任何 $n\in\mathbb{N}$, $|x_n|<M$ 成立, 就说数列 $\{x_n\}$ 是*有界数列*.

定理 1 a) *最终为常数的数列收敛.*

b) *数列极限的任何邻域, 包含着数列中除了有限多个数之外的所有项.*

c) *数列不能有两个不同的极限.*

d) *收敛数列必有界.*

◀ a) 如果当 $n>N$ 时, $x_n=A$, 那么对于点 A 的任何邻域 $V(A)$, 当 $n>N$ 时, $x_n\in V(A)$, 即 $\lim\limits_{n\to\infty}x_n=A$.

b) 此结论由数列极限的定义直接得到.

c) 这是本定理的重点. 设 $\lim\limits_{n\to\infty} x_n = A_1$ 且 $\lim\limits_{n\to\infty} x_n = A_2$. 如果 $A_1 \neq A_2$. 那么就能定出 A_1, A_2 的两个不相交的邻域 $V(A_1), V(A_2)$.

例如, 可以取 A_1, A_2 的 δ 邻域, 这里取 $\delta = \frac{1}{2}|A_1 - A_2|$. 据极限之定义, 求得 N_1 与 N_2, 使得

$$\forall n > N_1(x_n \in V(A_1)), \quad \forall n > N_2(x_n \in V(A_2)).$$

因此, 当 $n > \max\{N_1, N_2\}$ 时, 有 $x_n \in V(A_1) \cap V(A_2)$. 但因 $V(A_1) \cap V(A_2) = \varnothing$, 所以, 这是不可能的.

d) 设 $\lim\limits_{n\to\infty} x_n = A$. 在极限定义中, 令 $\varepsilon = 1$, 就有 N 使得 $\forall n > N(|x_n - A| < 1)$. 这就是说, 当 $n > N$ 时有 $|x_n| < |A| + 1$. 取

$$M > \max\{|x_1|, |x_2|, \cdots, |x_n|, |A| + 1\}$$

就得知 $\forall n > N(x_n < M)$. ►

b. 极限过程与算术运算

定义 6 设 $\{x_n\}, \{y_n\}$ 是两个数列, 那么分别称数列

$$\{(x_n + y_n)\}, \{(x_n y_n)\}, \left\{\left(\frac{x_n}{y_n}\right)\right\}$$

为它们的和、积与商(与函数的和、积与商的一般定义相符).

当然, 只有当 $y_n \neq 0, n \in \mathbb{N}$ 时, 商才有定义.

定理 2 *设 $\{x_n\}, \{y_n\}$ 是数列, 如果 $\lim\limits_{n\to\infty} x_n = A$, $\lim\limits_{n\to\infty} y_n = B$, 那么*

a) $\lim\limits_{n\to\infty} (x_n + y_n) = A + B$;

b) $\lim\limits_{n\to\infty} x_n \cdot y_n = A \cdot B$;

c) $\lim\limits_{n\to\infty} \frac{x_n}{y_n} = \frac{A}{B}$, *如果 $y_n \neq 0 (n = 1, 2, \cdots)$ 且 $B \neq 0$.*

◄ 我们要利用已经知道的 (见第二章 §2 第 4 段) 用近似值进行算术运算时, 所产生的绝对误差的估计式.

命 $|A - x_n| = \Delta(x_n), |B - y_n| = \Delta(y_n)$. 这时, 对于 a), 我们有

$$|(A + B) - (x_n + y_n)| \leqslant \Delta(x_n) + \Delta(y_n).$$

设给定 $\varepsilon > 0$. 因为 $\lim\limits_{n\to\infty} x_n = A$, 存在号码 N', 使得 $\forall n > N'\left(\Delta(x_n) < \frac{\varepsilon}{2}\right)$. 因为 $\lim\limits_{n\to\infty} y_n = B$, 存在号码 N'', 使得 $\forall n > N''\left(\Delta(y_n) < \frac{\varepsilon}{2}\right)$. 所以, 当 $n > N = \max\{N', N''\}$ 时, 就得到

$$|(A + B) - (x_n + y_n)| < \varepsilon.$$

据极限的定义, 断语 a) 得证.

b) 已知

$$|A \cdot B - x_n \cdot y_n| \leqslant |x_n|\Delta(y_n) + |y_n|\Delta(x_n) + \Delta(x_n) \cdot \Delta(y_n).$$

对于给定的 $\varepsilon > 0$, 存在着数 N' 与 N'', 使得

$$\forall n > N' \left(\Delta(x_n) < \min\left\{1, \frac{\varepsilon}{3(|B|+1)}\right\}\right),$$

$$\forall n > N'' \left(\Delta(y_n) < \min\left\{1, \frac{\varepsilon}{3(|B|+1)}\right\}\right).$$

所以, 当 $n > N = \max\{N', N''\}$ 时, 就有

$$|x_n| < |A| + \Delta(x_n) < |A| + 1,$$

$$|y_n| < |B| + \Delta(y_n) < |B| + 1,$$

$$\Delta(x_n) \cdot \Delta(y_n) < \min\left\{1, \frac{\varepsilon}{3}\right\} \cdot \min\left\{1, \frac{\varepsilon}{3}\right\} < \frac{\varepsilon}{3},$$

于是, 当 $n > N$ 时

$$|x_n|\Delta(y_n) < (|A|+1) \cdot \frac{\varepsilon}{3(|A|+1)} = \frac{\varepsilon}{3},$$

$$|y_n|\Delta(x_n) < (|B|+1) \cdot \frac{\varepsilon}{3(|B|+1)} = \frac{\varepsilon}{3},$$

$$\Delta(x_n) \cdot \Delta(y_n) < \frac{\varepsilon}{3},$$

因此, 当 $n > N$ 时, $|AB - x_n y_n| < \varepsilon$.

c) 利用估计式

$$\left|\frac{A}{B} - \frac{x_n}{y_n}\right| \leqslant \frac{|x_n|\Delta(y_n) + |y_n| \cdot \Delta(x_n)}{y_n^2} \cdot \frac{1}{1 - \delta(y_n)},$$

这里 $\delta(y_n) = \dfrac{\Delta(y_n)}{|y_n|}$.

对于给定的 $\varepsilon > 0$, 存在数 N' 及 N'', 使得

$$\forall n > N' \left(\Delta(x_n) < \min\left\{1, \frac{\varepsilon|B|}{8}\right\}\right),$$

$$\forall n > N'' \left(\Delta(y_n) < \min\left\{\frac{|B|}{4}, \frac{\varepsilon \cdot B^2}{16(|A|+1)}\right\}\right).$$

所以, 当 $n > \max\{N', N''\}$ 时, 得到

$$|x_n| < |A| + \Delta(x_n) < |A| + 1,$$

$$|y_n| > |B| - \Delta(y_n) > |B| - \frac{|B|}{4} > \frac{|B|}{2},$$

$$\frac{1}{|y_n|} < \frac{2}{|B|},$$

$$0 < \delta(y_n) = \frac{\Delta(y_n)}{|y_n|} < \frac{|B|/4}{|B|/2} = \frac{1}{2},$$

$$1 - \delta(y_n) > \frac{1}{2},$$

因此,

$$|x_n| \cdot \frac{1}{y_n^2}\Delta(y_n) < (|A| + 1) \cdot \frac{4}{B^2} \cdot \frac{\varepsilon B^2}{16(|A| + 1)} = \frac{\varepsilon}{4},$$

$$\left|\frac{1}{y_n}\right| \cdot \Delta(x_n) < \frac{2}{|B|} \cdot \frac{\varepsilon |B|}{8} = \frac{\varepsilon}{4},$$

$$0 < \frac{1}{1 - \delta(y_n)} < 2,$$

于是, 得: 当 $n > N$ 时, 有

$$\left|\frac{A}{B} - \frac{x_n}{y_n}\right| < \varepsilon. ▶$$

注 这样表述的定理也可用另外的、较少构造性的方法去证明, 大概, 这是读者从中学课本的分析初步就已经知道的. 后面讲到任意函数的极限时, 我们将重温这种方法, 而在这里, 在我们讨论数列的极限时, 我们想关注一下, 根据对算术运算结果的误差要求, 寻找参与运算的量的容许误差的问题.

c. 极限过程与不等式

定理 3 a) 设 $\{x_n\}, \{y_n\}$ 是两个收敛数列, 并且 $\lim\limits_{n\to\infty} x_n = A$, $\lim\limits_{n\to\infty} y_n = B$. 如果 $A < B$, 就有号码 $N \in \mathbb{N}$, 使得对于任何 $n > N$, 不等式 $x_n < y_n$ 成立.

b) 设 $\{x_n\}, \{y_n\}, \{z_n\}$ 是这样三个数列: 当 $n > N \in \mathbb{N}$ 时, $x_n \leqslant y_n \leqslant z_n$. 如果 $\{x_n\}$ 与 $\{z_n\}$ 收敛于同一极限, 那么数列 $\{y_n\}$ 也收敛于这个极限.

◀ a) 取一数 C 使 $A < C < B$. 根据极限的定义, 存在数 N' 与 N'', 使得当 $n > N'$ 时, $|x_n - A| < C - A$, 并且当 $n > N''$ 时, $|y_n - B| < B - C$, 这样, 当 $n > N = \max\{N', N''\}$ 时, 就得到

$$x_n < A + (C - A) = C = B - (B - C) < y_n.$$

b) 设 $\lim\limits_{n\to\infty} x_n = \lim\limits_{n\to\infty} z_n = A$. 对于 $\varepsilon > 0$, 找出 N', N'', 使得当 $n > N'$ 时, $A - \varepsilon < x_n$, 当 $n > N''$ 时, $z_n < A + \varepsilon$. 于是, 当 $n > N = \max\{N'N''\}$ 时, 得到 $A - \varepsilon < x_n \leqslant y_n \leqslant z_n < A + \varepsilon$, 或 $|y_n - A| < \varepsilon$. 因此 $A = \lim\limits_{n\to\infty} y_n$. ▶

推论 设 $\lim x_n = A, \lim y_n = B$. 如果存在号码 N, 使得当 $n > N$ 时.

a) $x_n > y_n$,那么 $A \geqslant B$;

b) $x_n \geqslant y_n$,那么 $A \geqslant B$;

c) $x_n > B$,那么 $A \geqslant B$;

d) $x_n \geqslant B$,那么 $A \geqslant B$.

◀ 用反证法, 从定理的第 a) 条直接推出 a) 和 b) 来. 第二、四断语是前两个断语在 $y_n \equiv B$ 时的特殊情况. ▶

请注意, 一个严格不等式, 在求极限后可能变成等式. 例如, 对于一切的 $n \in \mathbb{N}, \frac{1}{n} > 0$, 但 $\lim\limits_{n\to\infty} \frac{1}{n} = 0$.

3. 数列极限的存在问题

a. 柯西准则

定义 7 数列$\{x_n\}$叫做*基本列*(或*柯西列*), 如果对于任何数 $\varepsilon > 0$, 存在号码 $N \in \mathbb{N}$, 使得由 $n > N, m > N$ 推知 $|x_m - x_n| < \varepsilon$.①

定理 4 (数列收敛的柯西准则) *数列收敛的充要条件是它是基本列.*

◀设 $\lim\limits_{n\to\infty} x_n = A$. 对于 $\varepsilon > 0$, 求出号码 N, 使得当 $n > N$ 时, $|x_n - A| < \frac{\varepsilon}{2}$. 如果 $m, n > N$, 那么

$$|x_m - x_n| < |x_m - A| + |x_n - A| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

因此, 验证了收敛数列必是基本数列.

今设 $\{x_k\}$ 是基本列. 对于给定的 $\varepsilon > 0$, 存在号码 N 使得当 $m \geqslant N$ 及 $k \geqslant N$ 时, 有 $|x_m - x_k| < \frac{\varepsilon}{3}$, 把 m 固定成 $m = N$, 那么, 对一切 $k > N$ 有

$$x_N - \frac{\varepsilon}{3} < x_k < x_N + \frac{\varepsilon}{3}. \tag{1}$$

但因数列 $\{x_k\}$ 中足码不超过 N 的只有有限项, 因此, 我们证明了基本数列是有界数列.

现在, 对于一切 $n \in \mathbb{N}$, 令

$$a_n := \inf_{k\geqslant n} x_k, b_n := \sup_{k\geqslant n} x_k.$$

从这些定义明显地看到 $a_n \leqslant a_{n+1} \leqslant b_{n+1} \leqslant b_n$(因为集合变小时, 其下界不减小, 上界不增大). 根据闭区间套原理, 闭区间套 $[a_n, b_n]$ 有公共点 A.

① 柯西数列是波尔察诺引进的. 他没去处理实数的精确概念而试图证明基本列的收敛性. 柯西把区间套原理看成是显然的, 并据此证明了这个收敛性. 而区间套原理是后来的康托尔才找到了它的依据.

因为对于任何 $n \in \mathbb{N}$, 总有

$$a_n \leqslant A \leqslant b_n,$$

而且, 当 $k \geqslant n$ 时, 有

$$a_n = \inf_{k \geqslant n} x_k \leqslant x_k \leqslant \sup_{k \geqslant n} x_k = b_n,$$

所以, 当 $k \geqslant n$ 时, 得到

$$|A - x_k| \leqslant b_n - a_n \tag{2}$$

但由 (1) 式推知, 当 $n > N$ 时,

$$x_N - \frac{\varepsilon}{3} \leqslant \inf_{k \geqslant n} x_k = a_n \leqslant b_n = \sup_{k \geqslant n} x_k \leqslant x_N + \frac{\varepsilon}{3},$$

所以当 $n > N$ 时,

$$b_n - a_n \leqslant \frac{2\varepsilon}{3} < \varepsilon. \tag{3}$$

比较 (2) 与 (3) 式, 得到, 当 $k > N$ 时, 有

$$|A - x_k| < \varepsilon,$$

这就证明了 $\lim\limits_{k \to \infty} x_k = A$. ▶

例 8 数列 $(-1)^n, (n = 1, 2, \cdots)$ 没有极限, 因为它不是基本列. 虽然这是很明显的, 我们还是正式地给予验证. $\{x_n\}$ 是基本列的否命题就是:

$$\exists \varepsilon > 0, \forall N \in \mathbb{N} \exists n > N, \exists m > N (|x_m - x_n| \geqslant \varepsilon),$$

即存在 $\varepsilon > 0$, 使对于任何 $N \in \mathbb{N}$, 存在着大于 N 的 n 与 m, 使得 $|x_m - x_n| \geqslant \varepsilon$.

在本例中, 只要取 $\varepsilon = 1$ 即可. 这时, 对于任何 $N \in \mathbb{N}$, 将有

$$|x_{N+1} - x_{N+2}| = |1 - (-1)| = 2 > 1 = \varepsilon.$$

例 9 令

$$x_1 = 0.\alpha_1, x_2 = 0.\alpha_1\alpha_2, x_3 = 0.\alpha_1\alpha_2\alpha_3, \cdots, x_n = 0.\alpha_1\alpha_2\cdots\alpha_n, \cdots.$$

是一个有限二进小数的数列, 并且, 每项后面的项, 是把该项末尾补充一个数码 0 或 1 得来的. 我们来证明这样的数列一定收敛. 令 $m > n$, 并估计差值 $x_m - x_n$:

$$|x_m - x_n| = \left|\frac{\alpha_{n+1}}{2^{n+1}} + \cdots + \frac{\alpha_m}{2^m}\right| \leqslant \frac{1}{2^{n+1}} + \cdots + \frac{1}{2^m} = \frac{\left(\frac{1}{2}\right)^{n+1} - \left(\frac{1}{2}\right)^{m+1}}{1 - \frac{1}{2}} < \frac{1}{2^n}.$$

于是, 对于给定的 $\varepsilon > 0$, 取数 N, 使得 $\frac{1}{2^N} < \varepsilon$, 所以, 当 $m > n > N$ 时, 就得到估计值

$$|x_m - x_n| < \frac{1}{2^n} < \frac{1}{2^N} < \varepsilon,$$

这就证明了 $\{x_n\}$ 是基本数列.

例 10 讨论数列 $\{x_n\}$, 这里

$$x_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}.$$

因为对于任何 $n \in \mathbb{N}$,

$$|x_{2n} - x_n| = \frac{1}{n+1} + \cdots + \frac{1}{n+n} > n \cdot \frac{1}{2n} = \frac{1}{2}.$$

所以, 根据柯西准则, 这数列没有极限.

b. 单调数列的极限存在准则

定义 8 设有数列 $\{x_n\}$. 如果 $\forall n \in \mathbb{N}(x_n < x_{n+1})$, 就说 $\{x_n\}$ 是*递增列*; 如果 $\forall n \in \mathbb{N}(x_n \leqslant x_{n+1})$ 就称之为*不降列*; 如果 $\forall n \in \mathbb{N}(x_n \geqslant x_{n+1})$, 就称之为*不增列*; 如果 $\forall n \in \mathbb{N}, (x_n > x_{n+1})$ 就称之为*递降列*. 这四种数列, 都叫*单调数列*.

定义 9 称数列 $\{x_n\}$ 为*上有界列*, 如果存在着数 M, 使得 $\forall n \in \mathbb{N}(x_n < M)$.

类似地可定义*下有界数列*.

定理 5 (魏尔斯特拉斯) *不降数列有极限的充要条件是它上有界.*

◀ 因为在讨论数列的普遍性质时, 已经证明了任何收敛数列必定是有界数列, 所以, 只需证明条件是充分的.

据题设, 数列 $\{x_n\}$ 上有界, 因此它有上确界 $s = \sup\limits_{n\in\mathbb{N}} x_n$, 由上确界之定义, 对于任意的 $\varepsilon > 0$, 存在元素 $x_N \in \{x_n\}$, 使得 $s - \varepsilon < x_N \leqslant s$, 因为数列 $\{x_n\}$ 不降, 对于任何 $n > N$, 得到

$$s - \varepsilon < x_N \leqslant x_n \leqslant s, \text{即} |s - x_n| = s - x_n < \varepsilon.$$

这证明了 $\lim\limits_{n\to\infty} x_n = s$. ▶

当然, 对于下有界的不增数列, 也有类似的定理, 这时 $\lim\limits_{n\to\infty} x_n = \inf\limits_{n\in\mathbb{N}} x_n$.

注 不降 (不增) 数列的上 (下) 有界性, 显然与其有界性等价.

我们来讨论几个有用的例子.

例 11 当 $q > 1$ 时, $\lim\limits_{n\to\infty} \frac{n}{q^n} = 0$.

◀ 实际上, 如果 $x_n = \frac{n}{q^n}$, 那么 $x_{n+1} = \frac{n+1}{nq} x_n, n \in \mathbb{N}$. 因为

$$\lim_{n\to\infty} \frac{n+1}{nq} = \lim_{n\to\infty} \left(1+\frac{1}{n}\right)\frac{1}{q} = \lim_{n\to\infty} \left(1+\frac{1}{n}\right) \lim_{n\to\infty} \frac{1}{q} = 1 \cdot \frac{1}{q} = \frac{1}{q} < 1,$$

所以, 存在号码 N, 使得当 $n > N$ 时, $\frac{n+1}{nq} < 1$. 因此, 当 $n > N$ 时, $x_{n+1} < x_n$, 即在 x_N 这项以后, 我们的数列就单调下降. 因为由定义知道, 数列的有限多项不影响它的收敛性及极限值, 所以, 现在只需求出数列 $x_{N+1}, x_{N+2}, \cdots$ 的极限.

数列各项都是正的, 从而是下有界的, 这说明它有极限.

令 $x = \lim_{n\to\infty} x_n$. 由 $x_{n+1} = \frac{n+1}{nq} x_n$ 推知

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \left(\frac{n+1}{nq} x_n\right) = \lim_{n\to\infty} \frac{n+1}{nq} \lim_{n\to\infty} x_n = \frac{1}{q} x,$$

由此得到 $\left(1-\frac{1}{q}\right) x = 0$, 所以 $x = 0$. ▶

推论 1 $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

◀ 对于确定的 $\varepsilon > 0$, 据例 11 必能找到 $N \in \mathbb{N}$, 使得当 $n > N$ 时, $1 \leqslant n < (1+\varepsilon)^n$, 因此, 当 $n > N$ 时, 得到 $1 \leqslant \sqrt[n]{n} < 1+\varepsilon$. 这就是说 $\lim_{n\to\infty} \sqrt[n]{n} = 1$. ▶

推论 2 对于任何 $a > 0$, $\lim_{n\to\infty} \sqrt[n]{a} = 1$.

◀ 设 $a \geqslant 1$. 对于任意的 $\varepsilon > 0$, 存在着 $N \in \mathbb{N}$, 使得当 $n > N$ 时, $1 \leqslant a < (1+\varepsilon)^n$, 因此, 当 $n > N$ 时, $1 \leqslant \sqrt[n]{a} < 1+\varepsilon$, 即 $\lim_{n\to\infty} \sqrt[n]{a} = 1$.

如果 $0 < a < 1$. 那么 $1 < \frac{1}{a}$, 因而

$$\lim_{n\to\infty} \sqrt[n]{a} = \lim_{n\to\infty} \frac{1}{\sqrt[n]{\frac{1}{a}}} = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{\frac{1}{a}}} = 1. ▶$$

例 12 $\lim_{n\to\infty} \frac{q^n}{n!} = 0$, 这里 q 是任意实数, $n \in \mathbb{N}, n! := 1 \cdot 2 \cdot \cdots \cdot n$.

◀ 如果 $q = 0$, 结论是显然的. 再因 $\left|\frac{q^n}{n!}\right| = \frac{|q|^n}{n!}$, 所以只对 $q > 0$ 的情形证明就行了. 这时, 像前面一样, 我们看到 $x_{n+1} = \frac{q}{n+1} x_n$. 因为自然数集不是上有界的, 所以存在着号码 N, 使得当 $n > N$ 时, $0 < \frac{q}{n+1} < 1$, 因此, 当 $n > N$ 时, $x_{n+1} < x_n$, 再注意序列的项都是正的, 就保证了极限 $\lim_{n\to\infty} x_n = x$ 存在, 而这时有

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \frac{q}{n+1} x_n = \lim_{n\to\infty} \frac{q}{n+1} \lim_{n\to\infty} x_n = 0 \cdot x = 0. ▶$$

c. 数 e

例 13 证明极限 $\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^n$ 存在.

我们讨论的这个极限是这样的一个数，它在分析中处于特殊的地位，就像算术中的 1 和几何中的 π 那样，追随欧拉，我们把这极限记作 e. 我们还将因各种极不相同的理由、不止一次地回到这个极限上来.

先来验证下面的不等式：

$$(1+\alpha)^n \geqslant 1+n\alpha, \text{其中} n\in\mathbb{N}, \alpha>-1.$$

(有时把这个不等式叫做伯努利①不等式).

◀ $n=1$ 时，等式成立. 如果对 $n\in\mathbb{N}$，不等式成立，那么对 $n+1$ 不等式也成立，因为这时有

$$\begin{aligned}(1+\alpha)^{n+1} &= (1+\alpha)(1+\alpha)^n \geqslant (1+\alpha)(1+n\alpha)\\ &= 1+(n+1)\alpha+n\alpha^2 \geqslant 1+(n+1)\alpha.\end{aligned}$$

据归纳法原理，不等式对任何 $n\in\mathbb{N}$ 成立.

顺便指出，由上面的推证看出，当 $\alpha\neq 0$ 且 $n>1$ 时，成立严格的不等式. ▶

现在证明数列 $y_n=\left(1+\frac{1}{n}\right)^{n+1}$ 是递降列.

◀ 设 $n\geqslant 2$. 利用刚才证明的不等式，得到

$$\begin{aligned}\frac{y_{n-1}}{y_n} &= \frac{\left(1+\frac{1}{n-1}\right)^n}{\left(1+\frac{1}{n}\right)^{n+1}} = \frac{n^{2n}}{(n^2-1)^n}\cdot\frac{n}{n+1}\\ &= \left(1+\frac{1}{n^2-1}\right)^n\cdot\frac{n}{n+1}\\ &\geqslant \left(1+\frac{n}{n^2-1}\right)\cdot\frac{n}{n+1} > \left(1+\frac{1}{n}\right)\cdot\frac{n}{n+1}=1.\end{aligned}$$

因为是正项数列，所以极限 $\lim\limits_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+1}$ 存在. 而这时有

$$\begin{aligned}\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n &= \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+1}\left(1+\frac{1}{n}\right)^{-1}\\ &= \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+1}\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{-1} = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+1}.\end{aligned}$$ ▶

于是有

① 雅各布 伯努利 (Jacob Bernoulli) (1654—1705) —— 瑞士数学家. 著名的伯努利学者家族的代表 变分法与概率论的启蒙者.

定义 10

$$\boxed{e := \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n.}$$

d. 子列与数列的部分极限

定义 11 设 $x_1, x_2, \cdots, x_n, \cdots$ 是一数列, 而 $n_1 < n_2 < \cdots < n_k < \cdots$ 是自然数的一个递增列, 则称 $x_{n_1}, x_{n_2}, \cdots, x_{n_k}, \cdots$ 为数列 $\{x_n\}$ 的一个子列.

例如, 当把奇自然数 $1, 3, 5, \cdots$ 按其天然的顺序来取时, 就是数列 $1, 2, 3, \cdots$ 的一个子列. 但是数列 $3, 1, 5, 7, 9, \cdots$ 不是数列 $1, 2, 3, \cdots$ 的子列.

引理 1 (波尔察诺–魏尔斯特拉斯引理) *每个有界实数列含有收敛的子列.*

◀ 设 E 是有界数列$\{x_n\}$ 的值集. 如果 E 是有限集, 那么至少存在一点 $x \in E$, 及号码数列 $n_1 < n_2 < \cdots$ 使得

$$x_{n_1} = x_{n_2} = \cdots = x.$$

子数列 $\{x_{n_k}\}$ 是常数列, 所以它收敛.

如果 E 是无限集, 那么, 按波尔察诺–魏尔斯特拉斯原理, 它至少有一个极限点 x. 因为 x 是 E 的极限点, 必能选得一个 $n_1 \in \mathbb{N}$, 使得 $|x_{n_1} - x| < 1$, 如果已经选得了 $n_k \in \mathbb{N}$, 使得 $|x_{n_k} - x| < \frac{1}{k}$, 那么, 由于 x 是 E 的极限点, 必能选得 $n_{k+1} \in \mathbb{N}$, 使得 $n_k < n_{k+1}$ 且 $|x_{n_{k+1}} - x| < \frac{1}{k+1}$.

因为 $\lim\limits_{k\to\infty} \frac{1}{k} = 0$, 所以, 所作出的数列 $x_{n_1}, x_{n_2}, \cdots, x_{n_k}, \cdots$ 必收敛于 x. ▶

定义 12 给定数列 $\{x_n\}$. 如果对于每个数 c, 存在号码 $N \in \mathbb{N}$, 使得当 $n > N$ 时 $x_n > c$, 就说数列 $\{x_n\}$ *趋于正无穷*, 并记作 $x_n \to +\infty$.

现在把这个定义以及另外两个类似的定义, 用逻辑符号表示出来:

$$x_n \to +\infty := \forall c \in \mathbb{R},\ \exists N \in \mathbb{N}\ \forall n > N\ (c < x_n),$$
$$x_n \to -\infty := \forall c \in \mathbb{R},\ \exists N \in \mathbb{N}\ \forall n > N\ (x_n < c),$$
$$x_n \to \infty := \forall c \in \mathbb{R},\ \exists N \in \mathbb{N}\ \forall n > N\ (|c| < |x_n|).$$

在后面两种情形, 就分别说数列 $\{x_n\}$ *趋于负无穷*及数列 $\{x_n\}$ *趋于无穷*.

注意, 一个无穷数列可能是无界的, 但既不趋于正无穷, 又不趋于负无穷, 也不趋向无穷. 例如 $x_n = n^{(-1)^n}$.

趋于无穷的数列, 我们并不认为是收敛列.

容易看出, 有了这些定义, 可对上面证明的引理作些补充, 叙述稍有不同.

引理 2　从任何实数列中, 或者能选出一个收敛子数列, 或者能选出一个趋于无穷的子数列.

◀ 只要对无界数列来讨论就行了. 这时, 对于每个 $k \in \mathbb{N}$, 取 $n_k \in \mathbb{N}$, 使得 $|x_{n_k}| > k$, 并且 $n_k < n_{k+1}$, 于是得到了趋于无穷的子数列 $\{x_{n_k}\}$. ▶

设 $\{x_k\}$ 是任何实数列. 如果它下有界, 那么就能做出 $i_n = \inf\limits_{k \geqslant n} x_k$ 之数列 $\{i_n\}$(在柯西准则的证明中曾经看到过这个数列), 因为对于任何 $n \in \mathbb{N}$ 有 $i_n \leqslant i_{n+1}$, 所以数列 $\{i_n\}$ 或者具有有限的极限 $\lim\limits_{n\to\infty} i_n = l$, 或者 $i_n \to +\infty$.

定义 13　数 $l = \lim\limits_{n\to\infty} \inf\limits_{k \geqslant n} x_k$ 叫做数列 $\{x_k\}$ 的下极限, 并记作

$$\varliminf_{k\to\infty} x_k \quad \text{或} \quad \liminf_{k\to\infty} x_k.$$

如果 $i_n \to +\infty$, 就说数列 $\{x_k\}$ 的下极限等于正无穷, 并记作

$$\varliminf_{k\to\infty} x_k = +\infty \quad \text{或} \quad \liminf_{k\to\infty} x_k = +\infty.$$

如果数列 $\{x_k\}$ 无下界, 那么, 对于任何 $n \in \mathbb{N}$ 都有 $i_n = \inf\limits_{k \geqslant n} x_k = -\infty$. 这时就说数列的下极限等于负无穷, 并记作

$$\varliminf_{k\to\infty} x_k = -\infty \quad \text{或} \quad \liminf_{k\to\infty} x_k = -\infty.$$

于是, 考虑到上面列举的所有可能, 现在把数列$\{x_k\}$ 的下极限的定义简短地记作:

$$\boxed{\varliminf_{k\to\infty} x_k := \lim_{n\to\infty} \inf_{k \geqslant n} x_k.}$$

类似地, 讨论数列 $s_n = \sup\limits_{k \geqslant n} x_k$, 就得到数列 $\{x_k\}$ 的上极限的定义

定义 14

$$\boxed{\varlimsup_{k\to\infty} x_k := \lim_{n\to\infty} \sup_{k \geqslant n} x_k.}$$

举几个例子.

例 14

$$x_k = (-1)^k, k \in \mathbb{N}.$$

$$\varliminf_{k\to\infty} x_k = \lim_{n\to\infty} \inf_{k \geqslant n} x_k = \lim_{n\to\infty} \inf_{k \geqslant n} (-1)^k = \lim_{n\to\infty} (-1) = -1,$$

$$\varlimsup_{k\to\infty} x_k = \lim_{n\to\infty} \sup_{k \geqslant n} x_k = \lim_{n\to\infty} \sup_{k \geqslant n} (-1)^k = \lim_{n\to\infty} (+1) = 1.$$

例 15

$$x_k = k^{(-1)^k}, k \in \mathbf{N}.$$

$$\varliminf_{k\to\infty} k^{(-1)^k} = \lim_{n\to\infty} \inf_{k\geqslant n} k^{(-1)^k} = \lim_{n\to\infty} 0 = 0,$$

$$\varlimsup_{k\to\infty} k^{(-1)^k} = \lim_{n\to\infty} \sup_{k\geqslant n} k^{(-1)^k} = \lim_{n\to\infty} (+\infty) = +\infty.$$

例 16

$$x_k = k, k \in \mathbf{N}.$$

$$\varliminf_{k\to\infty} k = \lim_{n\to\infty} \inf_{k\geqslant n} k = \lim_{n\to\infty} n = +\infty.$$

$$\varlimsup_{k\to\infty} k = \lim_{n\to\infty} \sup_{k\geqslant n} k = \lim_{n\to\infty} (+\infty) = +\infty.$$

例 17

$$x_k = \frac{(-1)^k}{k}, k \in \mathbf{N}.$$

$$\varliminf_{k\to\infty} \frac{(-1)^k}{k} = \lim_{n\to\infty} \inf_{k\geqslant n} \frac{(-1)^k}{k}$$

$$= \lim_{n\to\infty} \left\{ \begin{array}{l} -\frac{1}{n}, \text{当} n = 2m+1 \\ -\frac{1}{n+1}, \text{当} n = 2m \end{array} \right\} = 0,$$

$$\varlimsup_{k\to\infty} \frac{(-1)^k}{k} = \lim_{n\to\infty} \sup_{k\geqslant n} \frac{(-1)^k}{k}$$

$$= \lim_{n\to\infty} \left\{ \begin{array}{l} \frac{1}{n}, \text{当} n = 2m \\ \frac{1}{n+1}, \text{当} n = 2m+1 \end{array} \right\} = 0.$$

例 18

$$x_k = -k^2, k \in \mathbf{N}.$$

$$\varliminf_{k\to\infty} (-k^2) = \lim_{n\to\infty} \inf_{k\geqslant n} (-k^2) = -\infty.$$

例 19

$$x_k = (-1)^k k, k \in \mathbf{N}.$$

$$\varliminf_{k\to\infty} (-1)^k k = \lim_{n\to\infty} \inf_{k\geqslant n} (-1)^k k = \lim_{n\to\infty} (-\infty) = -\infty,$$

$$\varlimsup_{k\to\infty} (-1)^k k = \lim_{n\to\infty} \sup_{k\geqslant n} (-1)^k k = \lim_{n\to\infty} (+\infty) = +\infty.$$

为了弄清楚数列的"上"极限与"下"极限这两个术语的来源，我们引进下面的

定义 15 如果数列 $\{x_n\}$ 有一子列收敛到某数 x(或趋于 $+\infty$, 或趋于 $-\infty$)，我们就说 x(或 $+\infty$, 或 $-\infty$) 是这个数列 $\{x_n\}$ 的一个部分极限.

命题 1 有界数列的下极限与上极限，分别是其部分极限中的最小者与最大者[①].

◀ 我们以下极限 $i = \varliminf\limits_{k\to\infty} x_k$ 为例来证明这个命题. 已经知道数列 $i_n = \inf\limits_{k\geqslant n} x_k$ 是不降列，并且 $\lim\limits_{n\to\infty} i_n = i \in \mathbb{R}$. 对于每个 $n \in \mathbb{R}$, 利用下确界的定义，据归纳法选出数 $k_n \in \mathbb{N}$，使得 $i_n \leqslant x_{k_n} < i_n + \frac{1}{n}$, 以及 $k_n < k_{n+1}$. 因为

$$\lim_{n\to\infty} i_n = \lim_{n\to\infty}\left(i_n + \frac{1}{n}\right) = i,$$

所以，根据极限的性质可知 $\lim\limits_{n\to\infty} x_{k_n} = i$, 这样，就证明了 i 是 $\{x_k\}$ 的一个部分极限. 它还是最小的部分极限，因为对于每个 $\varepsilon > 0$, 存在数 $n \in \mathbb{N}$, 使得 $i - \varepsilon < i_n$, 即当 $k \geqslant n$ 时，$i - \varepsilon < i_n = \inf\limits_{k\geqslant n} x_k \leqslant x_k$.

不等式 $i - \varepsilon < x_k$ 在 $k > n$ 时成立，这说明，数列 $\{x_k\}$ 的任何部分极限不能小于 $i - \varepsilon$. 但 ε 是任意的，所以这个部分极限也不能小于 i.

对于上极限，显然，可以类似地去证. ▶

我们指出，如果数列无下界，那么从其中可选出一个趋于 $-\infty$ 的子数列来，这时 $\varliminf\limits_{k\to\infty} x_k = -\infty$, 从而可以认为下极限仍是部分极限中最小的. 而这时的上极限可能是有限的，从而据上所证，它是部分极限中最大的；它也可能是无限. 如果 $\varlimsup\limits_{k\to\infty} x_k = +\infty$, 那么数列没有上界. 从而可以选出一个趋于 $+\infty$ 的子数列来. 如果 $\varlimsup\limits_{n\to\infty} x_k = -\infty$, 这种情况也是可能出现的，这表示 $\sup\limits_{k\geqslant n} x_k = s_n \to -\infty$, 也就是数列 $\{x_k\}$ 本身趋于 $-\infty$, 因为 $s_n \geqslant x_k$. 与此类似，如果 $\varliminf\limits_{k\to\infty} x_k = +\infty$, 那么 $x_k \to +\infty$.

总结上面的讨论，即得下面结论.

命题 1′ 任何数列的下极限，是它的部分极限中的最小者，而它的上极限，是它的部分极限中的最大者.

推论 1 数列有极限或趋于负无穷或趋于正无穷，其充要条件是其上、下极限重合.

---

① 同时，我们认为数 $x \in \mathbb{R}$ 与记号 $-\infty, +\infty$ 之间满足自然的关系 $-\infty < x < +\infty$.

◀ 当 $\underline{\lim}_{k\to\infty} x_k = \overline{\lim}_{k\to\infty} x_k = +\infty$, 或当 $\underline{\lim}_{k\to\infty} x_k = \overline{\lim}_{k\to\infty} x_k = -\infty$ 之情况, 上面已经讨论过了. 所以, 可以认为 $\underline{\lim}_{k\to\infty} x_k = \overline{\lim}_{k\to\infty} x_k = A \in \mathbb{R}$. 因为

$$i_n = \inf_{k\geqslant n} x_k \leqslant x_n \leqslant \sup_{k\geqslant n} x_k = s_n,$$

且根据题设有 $\lim_{n\to\infty} i_n = \lim_{n\to\infty} s_n = A$, 所以, 利用极限的性质, 有

$$\lim_{n\to\infty} x_n = A. ▶$$

推论 2 *数列收敛的充要条件是它的每个子数列收敛.*

◀ 子数列的下极限与上极限在原数列的下极限与上极限之间. 如果数列收敛, 那么, 它的下、上极限重合. 这时子列的下、上极限也重合, 由此推出它的收敛性, 并且所有子列有相同的极限.

逆命题是显然的, 因为数列本身就是一个子列. ▶

推论 3 *波尔察诺–魏尔斯特拉斯引理不论是狭义的还是广义的形式, 都可分别由命题 1 与命题 1′ 推出来.*

◀ 实际上, 如果数列 $\{x_k\}$ 有界, 那么, 点

$$i = \underline{\lim}_{k\to\infty} x_k 与 s = \overline{\lim}_{k\to\infty} x_k$$

都有限, 并且它们都是数列的部分极限. 只有当 $i = s$ 时, 数列有唯一的极限点; 当 $i < s$ 时, 就至少有两个了.

如果数列无界. 不论是哪边无界, 都存在着趋向于对应的那个无穷的子数列. ▶

结语 我们已完成了 (甚至有些超额) 本节开始时提出的三件事: 对数列的极限给出精确的定义; 证明极限的唯一性; 阐明极限运算与实数集结构之间的联系; 得到了数列收敛的判定准则.

现在我们来讨论数列的一种特殊的、经常遇到而且非常有用的形式 —— 级数.

4. 级数的初步知识

a. 级数的和与级数收敛性的柯西准则 设 $\{a_n\}$ 是实数列, 用 $\sum_{n=p}^{q} a_n$ 记和式 $a_p + a_{p+1} + \cdots + a_q (p \leqslant q)$. 我们现在对于由数列 $\{a_n\}$ 的项组成的表达式 $a_1 + a_2 + \cdots + a_n + \cdots$ 赋予确切的含义, 它隐含的意思是数列 $\{a_n\}$ 所有项之和.

定义 16 式子 $a_1 + a_2 + \cdots + a_n + \cdots$, 用记号 $\sum_{n=1}^{\infty} a_n$ 表示, 叫做**级数**或**无穷级数**(以区别于有限多个加项之和).

定义 17 把数列 $\{a_n\}$ 的元素看做是级数的元素, 并称之为级数的项; 把 a_n 叫做级数的第 n 项.

定义 18 称和 $s_n=\sum_{k=1}^{n}a_k$ 为级数的部分和; 为了指出它是多少项的和, 就称 s_n 为级数的第 n 部分和[①].

定义 19 如果级数的部分和组成的数列 $\{s_n\}$ 收敛, 就说级数收敛; 如果数列 $\{s_n\}$ 没有极限, 就说级数发散.

定义 20 如果部分和数列的极限存在: $\lim_{n\to\infty}s_n=s$, 就称它为级数的和.

今后我们就是在这个意义下理解式子.

$$\sum_{n=1}^{\infty}a_n=s$$

因为级数的收敛等价于其部分和数列 $\{s_n\}$ 的收敛, 所以把柯西准则应用于 $\{s_n\}$, 就立刻得到

定理 6 (级数收敛的柯西准则) *级数 $a_1+a_2+\cdots+a_n+\cdots$ 收敛的充要条件是, 对于任何 $\varepsilon>0$, 存在着 $N\in\mathbb{N}$ 使得当 $m\geqslant n>N$ 时,*

$$|a_n+\cdots+a_m|<\varepsilon.$$

推论 4 *如果在级数中只更换有限多项, 如果原来的级数收敛, 则新级数也收敛; 如果原来的级数发散, 则新级数也发散.*

◀ 为了证明这个推论, 只要认为柯西准则中出现的号码 N, 超过了所更换项的最大号码就行了. ▶

推论 5 *级数 $a_1+\cdots+a_n+\cdots$ 收敛的一个必要条件是当 $n\to\infty$ 时, 它的 (通) 项趋于零, 即必须*

$$\lim_{n\to\infty}a_n=0.$$

◀ 这只要在准则中令 $m=n$, 并利用数列极限的定义即知. ▶

另证: $a_n=s_n-s_{n-1}$. 既然已知 $\lim_{n\to\infty}s_n=s$, 于是,

$$\lim_{n\to\infty}a_n=\lim_{n\to\infty}(s_n-s_{n-1})=\lim_{n\to\infty}s_n-\lim_{n\to\infty}s_{n-1}=s-s=0.$$

例 20 级数 $1+q+q^2+\cdots+q^n+\cdots$ 通常叫做无穷几何级数(或等比级数) 之和. 现在讨论它的收敛性.

---

① 这样, 我们实际上是把级数理解成由关系 $\forall n\in\mathbb{N}\left(s_n=\sum_{k=1}^{n}a_k\right)$ 联系起来的数列对 $(\{a_n\},\{s_n\})$.

因为 $|q^n|=|q|^n$, 所以当 $|q|\geqslant 1$ 时 $|q^n|\geqslant 1$, 这时, 级数不满足收敛的必要条件.

今设 $|q|<1$. 这时

$$s_n = 1+q+\cdots+q^n = \frac{1-q^{n+1}}{1-q},$$

而 $\lim\limits_{n\to\infty} s_n = \dfrac{1}{1-q}$, 这是因为当 $|q|<1$ 时, $\lim\limits_{n\to\infty} q^{n+1}=0$.

因此, 级数 $\sum\limits_{n=1}^{\infty} q^n$ 收敛的充要条件是 $|q|<1$, 这时, 它的和是 $\dfrac{1}{1-q}$.

例 21 级数 $1+\frac{1}{2}+\cdots+\frac{1}{n}+\cdots$ 叫做*调和级数*, 因为这级数的每一项 (从第二项开始) 都是它相邻项的调和平均值 (参看本节末的习题 6).

级数的通项趋于零, 但其部分和

$$s_n = 1+\frac{1}{2}+\cdots+\frac{1}{n}$$

的数列发散, 这已在例 10 中用柯西准则证过, 也就是当 $n\to\infty$ 时 $s_n\to+\infty$.

所以调和级数是发散级数.

例 22 讨论下例.

级数 $1-1+1-\cdots+(-1)^{n+1}+\cdots$ 发散, 这可从它的部分和数列是 $1,0,1,0,\cdots$ 看出, 也可从级数的通项不趋于零看出.

如果添上括号, 并讨论这个新级数

$$(1-1)+(1-1)+\cdots$$

括号里的和是新级数的项, 这个新级数收敛, 并且其和为 0.

如果换一种方式加括号, 而讨论级数

$$1+((-1)+1)+(-1+1)+\cdots$$

得到的也是一个收敛级数, 它的和是 1.

如果把原来级数的项重排, 使等于 -1 的每个项, 向右移两位, 就得到级数,

$$1+1-1+1-1+1-\cdots$$

再添上括号就得到级数

$$(1+1)+(-1+1)+(-1+1)+\cdots$$

它的和是 2.

这种现象说明, 对于有限和所惯用的规则, 一般不能推广到级数上去.

然而毕竟有一类重要级数, 而后我们会证明, 能像有限和那样去处理它们. 这就是所谓的绝对收敛级数. 我们主要要讨论的正是这种级数.

b. 绝对收敛性; 比较定理及其推论

定义 21 如果级数 $\sum_{n=1}^{\infty}|a_n|$ 收敛, 就说级数 $\sum_{n=1}^{\infty}a_n$ 绝对收敛.

因为 $|a_n+\cdots+a_m|\leqslant|a_n|+\cdots+|a_m|$, 由柯西准则推知, 如果一个级数绝对收敛, 那么它必收敛.

一般来说, 它的逆命题不真, 即绝对收敛的要求, 比简单的收敛性的要求更强, 可以用例子来证明这件事.

例 23 级数

$$1-1+\frac{1}{2}-\frac{1}{2}+\frac{1}{3}-\frac{1}{3}+\cdots$$

的部分和是 $\frac{1}{n}$ 或零, 所以它收敛于零.

当把它的项换成绝对值时, 所得的级数

$$1+1+\frac{1}{2}+\frac{1}{2}+\frac{1}{3}+\frac{1}{3}+\cdots$$

发散, 这可像对调和级数那样, 由柯西准则推出:

$$\begin{aligned}&\left|\frac{1}{n+1}+\frac{1}{n+1}+\cdots+\frac{1}{n+n}+\frac{1}{n+n}\right|\\ =\ &2\left(\frac{1}{n+1}+\cdots+\frac{1}{n+n}\right)>2n\cdot\frac{1}{n+n}=1.\end{aligned}$$

为了回答级数是否绝对收敛的问题, 只要研究非负项级数的收敛性就行了. 对此, 我们有

定理 7 (非负项级数的判敛准则) *非负项级数收敛的充要条件是它的部分和的数列有上界.*

◀ 这从级数收敛的定义和不降数列收敛的准则即可得知. 这里的不降数列是级数的部分和数列 $s_1\leqslant s_2\leqslant\cdots\leqslant s_n\leqslant\cdots$. ▶

由这一准则可得到下面简单却非常实用的判断法.

定理 8 (比较定理) *设 $\sum_{n=1}^{\infty}a_n,\sum_{n=1}^{\infty}b_n$ 是两个非负项级数, 如果存在着号码 $N\in\mathbb{N}$, 使得当 $n>N$ 时 $a_n\leqslant b_n$, 那么, 当级数 $\sum_{n=1}^{\infty}b_n$ 收敛时, $\sum_{n=1}^{\infty}a_n$ 亦必收敛; 而当 $\sum_{n=1}^{\infty}a_n$ 发散时, $\sum_{n=1}^{\infty}b_n$ 亦必发散.*

◀ 因为级数的有限多项不影响级数的收敛性, 所以可以认为对一切 $n \in \mathbb{N}$ 都有 $a_n \leqslant b_n$ 而不失普遍性. 这时,

$$A_n = \sum_{k=1}^{n} a_k \leqslant \sum_{k=1}^{n} b_k = B_n.$$

如果级数 $\sum\limits_{n=1}^{\infty} b_n$ 收敛, 那么不降数列 $\{B_n\}$ 趋于极限 B. 这时, 对于任何 $n \in \mathbb{N}$, 有 $A_n \leqslant B_n \leqslant B$. 因此, 级数 $\sum\limits_{n=1}^{\infty} a_n$ 的部分和数列 $\{A_n\}$ 有界. 根据非负项级数的判敛准则知 $\sum\limits_{n=1}^{\infty} a_n$ 收敛.

定理的第二个断语, 由反证法即可从第一断语推出. ▶

例 24 因为当 $n \geqslant 2$ 时, 成立

$$\frac{1}{n(n+1)} < \frac{1}{n^2} < \frac{1}{(n-1)n}.$$

据比较定理可知, 级数 $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ 与 $\sum\limits_{n=1}^{\infty} \frac{1}{n(n+1)}$ 同时收敛或同时发散.

然而, 注意到 $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$, 从而, $\sum\limits_{k=1}^{n} \frac{1}{k(k+1)} = 1 - \frac{1}{n+1}$, 这后一级数能直接求和, 即 $\sum\limits_{n=1}^{\infty} \frac{1}{n(n+1)} = 1$. 因此, 级数 $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ 也收敛. 有趣的是 $\sum\limits_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$, 以后我们将证明这一点.

例 25 应注意, 比较定理只是关于非负项级数的. 实际上, 例如令 $a_n = -n$, $b_n = 0$, 则 $a_n < b_n$, 并且 $\sum\limits_{n=1}^{\infty} b_n$ 收敛. 但是 $\sum\limits_{n=1}^{\infty} a_n$ 发散.

推论 1 (魏尔斯特拉斯比较检验法) *设 $\sum\limits_{n=1}^{\infty} a_n$ 与 $\sum\limits_{n=1}^{\infty} b_n$ 是两个级数. 设存在着号码 N, 使得当 $n > N$ 时 $|a_n| \leqslant b_n$. 那么, 只要级数 $\sum\limits_{n=1}^{\infty} b_n$ 收敛, $\sum\limits_{n=1}^{\infty} a_n$ 就绝对收敛.*

◀ 实际上, 由比较定理知级数 $\sum\limits_{n=1}^{\infty} |a_n|$ 收敛, 这就是说 $\sum\limits_{n=1}^{\infty} a_n$ 绝对收敛. ▶

这个重要的有关绝对收敛的充分检验法, 时常简述为: *如果一级数的项* (按绝对值) *受控于收敛数项级数的项, 则原级数绝对收敛.*

例 26 级数 $\sum_{n=1}^{\infty}\frac{\sin n}{n^2}$ 绝对收敛. 因为 $\left|\frac{\sin n}{n^2}\right| \leqslant \frac{1}{n^2}$, 而级数 $\sum_{n=1}^{\infty}\frac{1}{n^2}$, 已在例 24 中看到, 是收敛的.

推论 2 (柯西根值检验法) 设 $\sum_{n=1}^{\infty}a_n$ 是给定的级数, 而

$$\alpha = \varlimsup_{n\to\infty}\sqrt[n]{|a_n|},$$

这时, 以下命题成立:

a) 如果 $\alpha < 1$, 则级数 $\sum_{n=1}^{\infty}a_n$ 绝对收敛;

b) 如果 $\alpha > 1$, 则级数 $\sum_{n=1}^{\infty}a_n$ 发散;

c) 在使 $\alpha = 1$ 的级数中, 既有绝对收敛的, 也有发散的.

◀ a) 如果 $\alpha < 1$, 就能选得一数 $q \in \mathbb{R}$, 使得 $\alpha < q < 1$. 固定数 q, 根据上极限之定义, 能求得号码 $N \in \mathbb{N}$, 使得当 $n > N$ 时, 有 $\sqrt[n]{|a_n|} < q$. 因此, 当 $n > N$ 时, 将有$|a_n| < q^n$, 又因当 $|q| < 1$ 时, $\sum_{n=1}^{\infty}q^n$ 收敛, 所以 (据比较定理或魏尔斯特拉斯比较检验法) 级数 $\sum_{n=1}^{\infty}a_n$ 绝对收敛.

b) 因为 α 是数列 $\{a_n\}$ 的部分极限 (参看命题 1), 找出使 $\lim_{k\to\infty}\sqrt[n_k]{|a_{n_k}|} = \alpha$ 的子数列 $\{a_{n_k}\}$. 当 $\alpha > 1$ 时, 可求得号码 $K \in \mathbb{N}$, 使当 $k > K$ 时, 有 $|a_{n_k}| > 1$. 因此级数 $\sum_{n=1}^{\infty}a_n$ 不满足收敛性的必要条件 $(a_n \to 0)$, 因而它发散.

c) 我们已知道级数 $\sum_{n=1}^{\infty}\frac{1}{n}$ 发散, 而级数 $\frac{1}{n^2}$ 收敛 $\left(\text{绝对收敛, 因为}\left|\frac{1}{n^2}\right| = \frac{1}{n^2}\right)$, 同时,

$$\varlimsup_{n\to\infty}\sqrt[n]{\frac{1}{n}} = \lim_{n\to\infty}\sqrt[n]{\frac{1}{n}} = \lim_{n\to\infty}\frac{1}{\sqrt[n]{n}} = 1,$$

且

$$\varlimsup_{n\to\infty}\sqrt[n]{\frac{1}{n^2}} = \lim_{n\to\infty}\sqrt[n]{\frac{1}{n^2}} = \lim_{n\to\infty}\left(\frac{1}{\sqrt[n]{n}}\right)^2 = 1. \blacktriangleright$$

例 27 讨论当 $x \in \mathbb{R}$ 取什么值时, 级数

$$\sum_{n=1}^{\infty}(2 + (-1)^n)^n x^n$$

收敛.

计算 $\alpha = \overline{\lim\limits_{n\to\infty}} \sqrt[n]{|(2+(-1)^n)^n x^n|} = |x| \overline{\lim\limits_{n\to\infty}} |2+(-1)^n| = 3|x|$. 因此, 当 $|x| < \dfrac{1}{3}$ 时级数收敛, 甚至是绝对收敛, 而当 $|x| > \dfrac{1}{3}$ 时, 级数发散. 在 $|x| = \dfrac{1}{3}$ 时, 需特别考虑. 但在本例中, 这是很容易的, 因为, 当 $|x| = \dfrac{1}{3}$ 时, 对于偶数的 n, 有 $(2+(-1)^{2k})^{2k} \cdot x^{2k} = 3^{2k} \cdot \left(\dfrac{1}{3}\right)^{2k} = 1$, 这说明级数不满足收敛的必要条件, 因而级数发散.

推论 3 (达朗贝尔[①]比值检验法) 设 $\sum\limits_{n=1}^{\infty} a_n$ 是一级数, 极限

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \alpha$$

存在, 这时, 以下命题成立:

a) 当 $\alpha < 1$ 时, 级数 $\sum\limits_{n=1}^{\infty} a_n$ 绝对收敛;

b) 当 $\alpha > 1$ 时, 级数 $\sum\limits_{n=1}^{\infty} a_n$ 发散;

c) 在使 $\alpha = 1$ 的级数中, 既有绝对收敛的也有发散的.

◀ a) 设 $\alpha < 1$, 于是有 q 使 $\alpha < q < 1$, 固定 q, 由极限的性质, 存在号码 $N \in \mathbb{N}$, 使得当 $n > N$ 时, 有 $\left|\dfrac{a_{n+1}}{a_n}\right| < q$. 因为改变级数的有限多项不会影响级数的收敛性质, 所以可以认为对于一切的 $n \in \mathbb{N}$ 有 $\left|\dfrac{a_{n+1}}{a_n}\right| < q$ 而不失普遍性.

因为

$$\left|\frac{a_{n+1}}{a_n}\right|\left|\frac{a_n}{a_{n-1}}\right| \cdot \ldots \cdot \left|\frac{a_2}{a_1}\right| = \left|\frac{a_{n+1}}{a_1}\right|,$$

所以 $|a_{n+1}| \leqslant |a_1| q^n$. 但级数 $\sum\limits_{n=1}^{\infty} a_1 q^n$ 收敛 $\left(\text{其和显然是}\dfrac{|a_1|}{1-q}\right)$, 所以级数 $\sum\limits_{n=1}^{\infty} a_n$ 绝对收敛.

b) 如果 $\alpha > 1$, 那么, 自某个号码 $N \in \mathbb{N}$ 开始, 只要 $n > N$ 就有 $\left|\dfrac{a_{n+1}}{a_n}\right| > 1$, 即 $|a_n| < |a_{n+1}|$. 因此, 级数 $\sum\limits_{n=1}^{\infty} a_n$ 不满足收敛的必要条件 $a_n \to 0$.

c) 像在柯西根值检验法中那样, 级数 $\sum\limits_{n=1}^{\infty} \dfrac{1}{n}$ 与 $\sum\limits_{n=1}^{\infty} \dfrac{1}{n^2}$ 可以做为这样的例子. ▶

例 28 现在来阐明 $x \in \mathbb{R}$ 取什么值时, 级数

$$\sum_{n=1}^{\infty} \frac{1}{n!} x^n$$

① 达朗贝尔(D' Alembert)(1717—1783)—— 法国学者, 但他首先是百科全书派的力学家.

收敛.

当 $x=0$ 时, 它显然收敛, 甚至是绝对收敛.

当 $x \neq 0$ 时, $\lim\limits_{n\to\infty}\left|\dfrac{a_{n+1}}{a_n}\right| = \lim\limits_{n\to\infty}\dfrac{|x|}{n+1} = 0.$

因此, 这级数对于一切 $x \in \mathbb{R}$ 都绝对收敛.

最后, 我们来讨论一类更特殊的、但经常遇到的级数: 其项组成单调数列的级数. 对于这样的级数, 有以下充要收敛检验法.

命题 2 (柯西) 如果 $a_1 \geqslant a_2 \geqslant \cdots \geqslant 0$, 那么, 级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛的充要条件是级数

$$\sum_{k=0}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + 8a_8 + \cdots$$

收敛.

◀ 因为

$$\begin{aligned}
a_2 &\leqslant a_2 \leqslant a_1,\\
2a_4 &\leqslant a_3 + a_4 \leqslant 2a_2,\\
4a_8 &\leqslant a_5 + a_6 + a_7 + a_8 \leqslant 4a_4,\\
&\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\
2^n a_{2^{n+1}} &\leqslant a_{2^n+1} + a_{2^n+2} + \cdots + a_{2^{n+1}} \leqslant 2^n a_{2^n},
\end{aligned}$$

将这些不等式相加就得到

$$\frac{1}{2}(s_{n+1} - a_1) \leqslant A_{2^{n+1}} - a_1 \leqslant s_n,$$

其中 $A_k = a_1 + \cdots + a_k, s_n = a_1 + 2a_2 + \cdots + 2^n a_{2^n}$ 是所讨论级数的部分和. 数列 $\{A_k\}$ 与 $\{s_n\}$ 都是不降的, 所以由得到的这个不等式可以推知, 它们或者同时有界, 或者同时无上界. 由此根据非负项级数的判敛准则可以推出, 这两个级数或者同时收敛, 或者同时发散. ▶

由此得到有用的

推论 当 $p>1$ 时级数 $\sum\limits_{n=1}^{\infty}\dfrac{1}{n^p}$ 收敛; 而当 $p \leqslant 1$ 时发散①.

◀ 如果 $p \geqslant 0$, 那么由命题 2, 此级数与级数

$$\sum_{k=0}^{\infty} 2^k \frac{1}{(2^k)^p} = \sum_{k=0}^{\infty} (2^{1-p})^k$$

① 在本书中, 我们暂时只对有理数 p 定义了 n^p, 所以读者也暂时将此命题理解成是对于使 n^p 有意义的 p 来说的.

同时收敛或同时发散，但后者收敛的充要条件是数 $q = 2^{1-p} < 1$，即 $p > 1$.

如果 $p \leqslant 0$，那么这时级数 $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ 的所有项都大于或等于 1，因此，它发散. ►

这个推论的重要性在于当讨论级数的收敛性时，时常用级数 $\sum\limits_{n=1}^{\infty} \frac{1}{n^p}$ 来作基本的对比级数.

c. 作为级数的和的数 e 在行将结束对级数的讨论时，我们再一次回到数 e 上来，并将得到一个为计算 e 提供了相当方便的方法的级数.

在把 $\left(1+\frac{1}{n}\right)^n$ 展开时要用到牛顿二项公式. 如果有人在中学里没有学过这个公式，也没有解答过第二章 §2 的习题 1.g，可以越过现在对于数 e 这段补充，这并不会接不上头. 等学了泰勒公式之后，再回到这里来，因为牛顿二项公式是泰勒公式的一个特殊情况.

我们已知，$e = \lim\limits_{n\to\infty} \left(1+\frac{1}{n}\right)^n$.

据牛顿二项公式

$$\begin{aligned}\left(1+\frac{1}{n}\right)^n &= 1 + \frac{n}{1!}\cdot\frac{1}{n} + \frac{n(n-1)}{2!}\cdot\frac{1}{n^2} + \cdots \\ &\quad + \frac{n(n-1)\cdot\cdots\cdot(n-k+1)}{k!}\cdot\frac{1}{n^k} + \cdots + \frac{1}{n^n} \\ &= 1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \cdots + \frac{1}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdot\cdots\cdot \\ &\quad \left(1-\frac{k-1}{n}\right) + \cdots + \frac{1}{n!}\left(1-\frac{1}{n}\right)\cdot\cdots\cdot\left(1-\frac{n-1}{n}\right).\end{aligned}$$

令 $\left(1+\frac{1}{n}\right)^n = e_n$，且 $1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!} = s_n$. 于是，$e_n < s_n (n = 1, 2, \cdots)$.

另一方面，对于任何固定的 k 及 $n \geqslant k$，仍从上面的展开式看到

$$1 + 1 + \frac{1}{2!}\left(1-\frac{1}{n}\right) + \cdots + \frac{1}{k!}\left(1-\frac{1}{n}\right)\cdot\cdots\cdot\left(1-\frac{k-1}{n}\right) < e_n.$$

当 $n \to \infty$ 时，此不等式左端趋于 s_k，而右端趋于 e，因此，现在推出了 $s_k \leqslant e$ 对一切 $k \in \mathbb{N}$ 成立.

但这时由关系式

$$e_n < s_n \leqslant e$$

令 $n \to \infty$ 就得到 $\lim\limits_{n\to\infty} s_n = e$.

根据级数和的定义，就能写出

$$\boxed{e = 1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + \cdots.}$$

e 的这个表示式, 对于计算 e 的值来说, 是完全合用的.

估计差值 $e - s_n$:

$$\begin{aligned} 0 < e - s_n &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \cdots \\ &= \frac{1}{(n+1)!}\left[1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \cdots\right] \\ &< \frac{1}{(n+1)!}\left[1 + \frac{1}{n+2} + \frac{1}{(n+2)^2} + \cdots\right] \\ &= \frac{1}{(n+1)!} \cdot \frac{1}{1 - \frac{1}{n+2}} = \frac{n+2}{n!(n+1)^2} < \frac{1}{n!n}. \end{aligned}$$

因此, 当用数 s_n 去近似代替数 e 时, 为了使绝对误差, 例如不超过 10^{-3}, 只要 $\frac{1}{n!n} < \frac{1}{1000}$ 就行了. s_6 已能满足这个条件.

e 的前若干 (十进) 位是

$$e = 2.7182818284590\cdots.$$

所得到的差 $e - s_n$ 的估计式, 可以写成等式的形式:

$$e = s_n + \frac{\theta_n}{n!n}, \text{这里} 0 < \theta_n < 1.$$

从数 e 的这种表示法可以直接得出它的无理性来. 事实上, 如果 $e = \frac{p}{q}, p, q \in \mathbb{N}$, 那么 $q!e$ 应是整数, 并且

$$q!e = q!\left(s_q + \frac{\theta_q}{q!q}\right) = q! + \frac{q!}{1!} + \frac{q!}{2!} + \cdots + \frac{q!}{q!} + \frac{\theta_q}{q},$$

所以 $\frac{\theta_q}{q}$ 也应是整数, 而这是不可能的.

提请读者注意, 数 e 不只是无理数, 它还是超越数.

练　习

1. 试证: 数 $x \in \mathbb{R}$ 是有理数的充要条件是, 它在任何 q 进位记数法中是循环的, 即从某一位开始, 就由一组数码重复循环组成.
2. 皮球从高为 h 之处落下时, 反跳起来的高度是 qh, 这里 q 是常数系数, $0 < q < 1$. 求它完全落地而不再跳起来所需的时间, 以及直到这个时刻前, 皮球所经过的路程.

3. 圆周上有一固定点, 把圆周转动 n 个弧度, n 取一切自然数, 就得到圆周上的许多点. 试求这样得到的点集的所有极限点.

4. 式子

$$n_1 + \cfrac{1}{n_2 + \cfrac{1}{n_3 + \ddots + \cfrac{1}{n_{k-1} + \cfrac{1}{n_k}}}}$$

叫做有限链式分数或叫有限连分数, 其中 $n_i \in \mathbf{N}$; 又称

$$n_1 + \cfrac{1}{n_2 + \cfrac{1}{n_3 + \ddots}}$$

为无穷连分数. 把一个连分数自某个链开始以后的链都去掉所得的分数, 叫做这个连分数的一个近似分数. 无穷连分数的近似分数数列的极限, 就作为这个无穷连分数的值

试证:

a) 每个有理数 $\frac{m}{n}$, $m, n \in \mathbf{N}$, 能唯一地展成连分数:

$$\frac{m}{n} = q_1 + \cfrac{1}{q_2 + \ddots + \cfrac{1}{q_{n-1} + \cfrac{1}{q_n}}}$$

提示 数 $q_1, \cdots, q_n$ 叫做不完全商, 它们由欧几里得辗转除法

$$\begin{aligned} m &= n \cdot q_1 + r_1 \\ n &= r_1 q_2 + r_2 \\ r_1 &= r_2 q_3 + r_3 \\ &\cdots\cdots\cdots\cdots \end{aligned}$$

的另一种写法 $\frac{m}{n} = q_1 + \frac{1}{n/r_1} = q_1 + \frac{1}{q_2 + r_1/r_2} = \cdots$ 得到.

b) 近似分数 $R_1 = q_1, R_2 = q_1 + \frac{1}{q_2}, \cdots$ 满足不等式

$$R_1 < R_3 < \cdots < R_{2k-1} < \frac{m}{n} < R_{2k} < R_{2k-2} < \cdots < R_2.$$

c) 近似分数 R_k 的分子 P_k 与分母 Q_k, 按以下的规律建立:

$$P_k = P_{k-1}q_k + P_{k-2}, P_2 = q_1 q_2, P_1 = q_1.$$

$$Q_k = Q_{k-1}q_k + Q_{k-2}, Q_2 = q_2, Q_1 = 1.$$

d) 两相邻近似分数之差按以下公式计算:

$$R_k - R_{k-1} = \frac{(-1)^k}{Q_k Q_{k-1}} \quad (k > 1).$$

e) 每个无穷连分数有确定的值.

f) 无穷连分数的值是无理数.

g)
$$\frac{1+\sqrt{5}}{2} = 1 + \cfrac{1}{1 + \cfrac{1}{1+ \ddots}}.$$

h) 做为 g) 中近似分数的分母得到的斐波纳契 (Fibonacci) 数 $1, 1, 2, 3, 5, 8, \cdots$(即 $u_n = u_{n-1} + u_{n-2}$ 且 $u_1 = u_2 = 1$), 可用公式

$$u_n = \frac{1}{\sqrt{5}}\left[\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n\right]$$

给出.

i) 上边 g) 中的近似分数 $R_k = \frac{P_k}{Q_k}$ 满足

$$\left|\frac{1+\sqrt{5}}{2} - \frac{P_k}{Q_k}\right| > \frac{1}{Q_k^2\sqrt{5}}$$

试比较它与第 II 章, §2, 习题 11 的断言.

5. 试证:

a) 当 $n \geqslant 2$ 时, 以下等式成立:

$$1 + \frac{1}{1!} + \frac{1}{2!} + \cdots + \frac{1}{n!} + \frac{1}{n!n} = 3 - \frac{1}{1\cdot 2\cdot 2!} - \cdots - \frac{1}{(n-1)n\cdot n!};$$

b) $e = 3 - \sum\limits_{n=0}^{\infty} \frac{1}{(n+1)(n+2)(n+2)!}$;

c) 当近似地计算 e 时, 用公式

$$e \approx 1 + \frac{1}{1!} + \cdots + \frac{1}{n!} + \frac{1}{n!n}$$

比用原来的公式

$$e \approx 1 + \frac{1}{1!} + \cdots + \frac{1}{n!}$$

好得多(作出误差估计, 作一个计算并与第 122 页上给出的 e 的值进行比较).

6. 如果 a 与 b 是正数, 而 p 是任意实数, 就称

$$S_p(a, b) = \left(\frac{a^p + b^p}{2}\right)^{\frac{1}{p}}$$

为 a 与 b 的 p次平均值.

特别当 $p = 1$ 时得到算术平均值; $p = 2$ 时得到二次平均值或均方值; 当 $p = -1$ 时, 得到调和平均值.

a) 试证任何次的平均值 $S_p(a,b)$ 界于 a,b 之间.

b) 求平均值列

$$\{S_n(a,b)\},\{S_{-n}(a,b)\}.$$

的极限.

7. 试证: 如果 $a>0$, 那么数列

$$x_{n+1}=\frac{1}{2}\left(x_n+\frac{a}{x_n}\right)$$

收敛于 a 的算术平方根, 其中 $x_1\neq 0$.

估计收敛速度, 即估计与 n 有关的绝对误差 $|\Delta_n|=|x_n-a|$.

8. 试证

a) $S_0(n)=1^0+\cdots+n^0=n$,

$$S_1(n)=1^1+\cdots+n^1=\frac{n(n+1)}{2}=\frac{1}{2}n^2+\frac{1}{2}n,$$
$$S_2(n)=1^2+\cdots+n^2=\frac{n(n+1)(2n+1)}{6}=\frac{1}{3}n^3+\frac{1}{2}n^2+\frac{1}{6}n,$$
$$S_3(n)=1^3+\cdots+n^3=\frac{n^2(n+1)^2}{4}=\frac{1}{4}n^4+\frac{1}{2}n^3+\frac{1}{4}n^2.$$

如果记 $S_k(n)=1^k+\cdots+n^k, k\in\mathbb{N}$, 那么

$$S_k(n)=a_{k+1}n^{k+1}+\cdots+a_1n+a_0,$$

它是 n 的 $k+1$ 次多项式.

b) $\lim\limits_{n\to\infty}\dfrac{S_k(n)}{n^{k+1}}=\dfrac{1}{k+1}$.

§2. 函数的极限

1. 定义和例子 设 E 是实数集 $\mathbb{R}$ 的一个子集, a 是 E 的一个极限点. 设 $f:E\to\mathbb{R}$ 是定义在 E 上的一个实值函数.

我们来描述当点 $x\in E$ 接近于 a 时, 函数 f 的值 $f(x)$ 接近于某数 A 的过程: 这时自然称数 A 为当 x 趋于 a 时函数 f 的极限值或极限.

定义 1 *设有函数 $f:E\to\mathbb{R}, A\in\mathbb{R}$. 如果对于任何 $\varepsilon>0$, 存在一数 $\delta>0$, 使得对于满足 $0<|x-a|<\delta$ 的任何 $x\in E$, 关系*

$$|f(x)-A|<\varepsilon$$

成立, 我们就(跟随柯西)*说当 x 趋于 a 时, 函数 f 趋于 A, 或说, A 是函数 f 当 x 趋于 a 时的极限.*

把所说的条件用逻辑符号写出来就是

$$\forall\varepsilon>0\exists\delta>0\forall x\in E(0<|x-a|<\delta\Rightarrow|f(x)-A|<\varepsilon).$$

如果当 x 沿集合 E 趋于 a 时 A 是函数 $f(x)$ 的极限, 就记成

$$f(x) \to A, \text{当} x \to a, x \in E, \text{或} \lim_{x \to a, x \in E} f(x) = A.$$

我们还常把 $x \to a, x \in E$ 写的更简单些:$E \ni x \to a$, 同时, 把 $\lim\limits_{x \to a, x \in E} f(x)$ 写成 $\lim\limits_{E \ni x \to a} f(x)$.

例 1 设 $E = \mathbb{R} \backslash 0, f(x) = x \sin \frac{1}{x}$. 验证

$$\lim_{E \ni x \to 0} x \sin \frac{1}{x} = 0.$$

实际上, 对给定的 $\varepsilon > 0$, 取 $\delta = \varepsilon$, 这时, 当 $0 < |x| < \delta = \varepsilon$ 时, 注意到 $\left| x \sin \frac{1}{x} \right| \leqslant |x|$, 就得 $\left| x \sin \frac{1}{x} \right| < \varepsilon$.

由这个例题, 顺便看到, 当 $E \ni x \to a$ 时, 即使函数 $f: E \to \mathbb{R}$ 在 a 点没有定义, 它也能有极限. 在极限计算中经常碰到的正是这种情况. 我们回过来想一想, 在用不等式 $0 < |x - a|$ 的形式来定义极限时, 我们就已经注意到这一点了.

大家都知道, 一点 $a \in \mathbb{R}$ 的邻域是包含这个点的一个任意的开区间.

定义 2 一点的邻域, 去掉这个点本身, 叫做这个点的*去心邻域*.

若$U(a)$表示点a的一个邻域, 我们就用$\mathring{U}(a)$表示它的去心邻域.

集合

$$U_E(a) := E \cap U(a),$$
$$\mathring{U}_E(a) := E \cap \mathring{U}(a)$$

分别叫做*点 a 在集合 E 中的邻域与去心邻域*.

如果 a 是 E 的极限点, 那么, 不论对 a 的哪个邻域 $U(a)$, 总有 $\mathring{U}_E(a) \neq \varnothing$.

如果用 $\mathring{U}_E^\delta(a)$与 $V_{\mathbb{R}}^\varepsilon(A)$ 这种累赘的记号表示点 a 在集合 E 中的去心 δ 邻域与 A 点在 $\mathbb{R}$ 中的 ε 邻域, 上面引进的函数的柯西 "$\varepsilon - \delta$ 定义" 可以改写成:

$$\left(\lim_{E \ni x \to a} f(x) = A \right) := \forall V_{\mathbb{R}}^\varepsilon(A) \exists \mathring{U}_E^\delta(a) \left(f(\mathring{U}_E^\delta(a)) \subset V_{\mathbb{R}}^\varepsilon(A) \right)$$

的形式.

这个写法是说: 如果对于点 A 的任何 ε 邻域 $V_{\mathbb{R}}^\varepsilon(A)$, 能找到点 a 在集合 E 中的 一个去心 δ 邻域 $\mathring{U}_E^\delta(a)$, 使得它在 $f: E \to \mathbb{R}$ 下的像 $f(\mathring{U}_E^\delta(a))$ 完全包含在邻域 $V_{\mathbb{R}}^\varepsilon(A)$ 中, 那么, A 就是当 x 沿集合 E 趋于 a 时, 函数 $f: E \to \mathbb{R}$ 的极限.

考虑到在数轴上一点的任何邻域中, 必含这一点的一个对称邻域 (δ 邻域), 我们现在得到极限定义的如下形式的写法, 并把它作为基本的写法.

定义 3

$$\left(\lim_{E\ni x\to a} f(x) = A\right) := \forall V_{\mathbb{R}}(A)\exists \mathring{U}_E(a)\left(f(\mathring{U}_E(a)) \subset V_{\mathbb{R}}(A)\right).$$

因此, 如果对于点 A 的任何邻域, 存在点 a 在集合 E 中的去心邻域, 使得它在映射 $f: E \to \mathbb{R}$ 下的像含于 A 的给定的这个邻域中, 就说数 A 是函数 $f: E \to \mathbb{R}$ 当 x 沿集合 E 趋于 (E 的极限) 点 a 时的极限.

我们已经引进了函数极限定义的几种说法. 我们知道, 对于数值函数, 即 $a, A \in \mathbb{R}$ 时, 它们是等价的. 同时, 为了不同的目的, 有时是这一种方便, 有时是那一种方便. 例如, 在估算函数值时, 以原来的形式较合适, 它指出了要想使得 $f(x)$ 对 A 的偏差不超过给定的数量时, x 对 a 所能容许的偏差是多少. 而当把极限概念推广到更一般的函数, 也就是函数的定义域不是数集时, 用最后这种定义的方式却是最合适的. 从这种定义顺便看到, 只要告诉我们 X 和 Y 中点的邻域是什么, 换句话说, 就是如果在 X 和 Y 中给定了拓扑, 我们就能定义映射 $f: X \to Y$ 的极限概念.

再考察几个说明基本定义的例子.

例 2 设在整个数轴上定义符号函数[①]

$$\operatorname{sgn} x = \begin{cases} 1, & 当\ x > 0, \\ 0, & 当\ x = 0, \\ -1, & 当\ x < 0, \end{cases}$$

今证, 当 x 趋于 0 时它没有极限.

也就是要证

$$\forall A \in \mathbb{R}\exists V(A)\forall \mathring{U}(0)\exists x \in \mathring{U}(0)(f(x) \notin V(A)),$$

亦即, 无论我取怎样的 A(妄想它是 $\operatorname{sgn} x$ 当 $x \to 0$ 时的极限), 都存在点 A 的邻域 $V(A)$, 使对点 0 的无论怎样 (小) 的去心邻域 $\mathring{U}(0)$, 至少存在一点 $x \in \mathring{U}(0)$, 使得函数 $\operatorname{sgn} x$ 在 x 的值不在 $V(A)$ 内.

因为函数 $\operatorname{sgn} x$ 只取 $-1, 0, 1$ 这三个值, 所以如果 A 不是其中之一, 它就不可能是函数的极限值, 因为这时它将有不含此三数中任何一个的邻域 $V(A)$.

如果 $A \in \{-1, 0, 1\}$, 则对 $\varepsilon = \dfrac{1}{2}$ 取 A 的 ε 邻域作为 $V(A)$. 点 1 和 -1 分明不能同时都落入这个邻域. 然而, 无论取点 0 的哪个去心邻域, 它都既含负数又含正数, 即在某些点上 $f(x) = -1$, 而在另外某些点上 $f(x) = 1$.

因此, 能找到点 $x \in \mathring{U}(0)$, 使 $f(x) \notin V(A)$.

我们约定, 如果函数 $f: E \to \mathbb{R}$ 在某点 $a \in \mathbb{R}$ 的整个去心邻域上有定义, 即当 $\mathring{U}_E(a) = \mathring{U}_{\mathbb{R}}(a) = \mathring{U}(a)$ 时, 将代替 $E \ni x \to a$ 使用比较简单的记法 $x \to a$.

[①] Signum(拉丁文)—— 符号

例 3　试证 $\lim\limits_{x\to 0}|\operatorname{sgn} x| = 1$.

实际上, 当 $x \in \mathbb{R}\backslash 0$ 时, $|\operatorname{sgn} x| = 1$, 即它在 0 点的任何去心邻域 $\mathring{U}(0)$ 中是常数 1, 因此, 对于点 1 的任何邻域 $V(1)$ 得到 $f(\mathring{U}(0)) = 1 \in V(1)$.

请注意, 虽然函数 $|\operatorname{sgn} x|$ 在点 0 已有定义 $|\operatorname{sgn} 0| = 0$, 但此值对函数的极限并无任何影响.

因此, 不要把函数在 a 点的值 $f(a)$ 与当 x 趋于 a 时的极限 $\lim\limits_{x\to a} f(x)$ 混为一谈.

设 $\mathbb{R}_-$ 与 $\mathbb{R}_+$ 分别表示负数集与正数集.

例 4　在例 2 中, 我们已知 $\lim\limits_{\mathbb{R}\ni x\to 0} \operatorname{sgn} x$ 不存在. 但是, 函数 sgn 在 $\mathbb{R}_-$ 上的限制 $\operatorname{sgn}|_{\mathbb{R}_-}$ 是等于 -1 的常数函数, 而 $\operatorname{sgn}|_{\mathbb{R}_+}$ 是等于 1 的常数函数. 可以像例 3 那样证明

$$\lim_{\mathbb{R}_-\ni x\to 0} \operatorname{sgn} x = -1, \lim_{\mathbb{R}_+\ni x\to 0} \operatorname{sgn} x = 1.$$

即同一个函数, 当把它限制在不同的子集上时, 在同一点处可有不同极限, 而原来的函数在这一点没有极限. 例 2 就是这样的.

例 5　把例 2 的思想推广开来, 可以类似地证明函数 $\sin\frac{1}{x}$ 当 $x \to 0$ 时没有极限.

实际上, 在点 0 的任何去心邻域 $\mathring{U}(0)$ 中, 总存在形如

$$\frac{1}{-\frac{\pi}{2} + 2\pi n} \quad 与 \quad \frac{1}{\frac{\pi}{2} + 2\pi n}$$

的点, 其中 $n \in \mathbb{N}$, 在这些点上, 函数分别取值 -1 与 $+1$. 但是当 $\varepsilon < 1$ 时, 这两个数不能同时含于点 $A \in \mathbb{R}$ 的 ε 邻域 $V(A)$ 内. 这就是说, 任何 $A \in \mathbb{R}$ 不能是这函数在 $x \to 0$ 时的极限.

例 6　如果

$$E_- = \left\{x \in \mathbb{R} \mid x = \frac{1}{-\frac{\pi}{2} + 2\pi n}, n \in \mathbb{N}\right\},$$

和

$$E_+ = \left\{x \in \mathbb{R} \mid x \in \frac{1}{\frac{\pi}{2} + 2\pi n}, n \in \mathbb{N}\right\}.$$

与例 4 类似, 得到

$$\lim_{E_-\ni x\to 0} \sin\frac{1}{x} = -1 \quad 及 \quad \lim_{E_+\ni x\to 0} \sin\frac{1}{x} = 1.$$

上节讨论的数列极限概念与这里引入的任意数值函数的极限概念有密切的联系. 以下命题反映了这种联系.

命题 1[①] 关系 $\lim\limits_{E\ni x\to a} f(x) = A$ 成立当且仅当对于任何收敛于 a 的点列 $\{x_n\}, x_n \in E\backslash a$, 数列 $\{f(x_n)\}$ 收敛于 A.

◀ 由定义立刻得到

$$(\lim_{E\ni x\to a} f(x) = A) \Rightarrow (\lim_{n\to\infty} f(x_n) = A),$$

实际上, 如果 $\lim\limits_{E\ni x\to a} f(x) = A$, 那么对于点 A 的任何邻域 $V(A)$, 存在 a 在 E 中的去心邻域 $\mathring{U}_E(a)$, 使得对于 $x \in \mathring{U}_E(a)$ 有 $f(x) \in V(A)$. 如果集合 $E \setminus a$ 的点列 $\{x_n\}$ 收敛于 a, 则存在号码 N, 使得当 $n > N$ 时, $x_n \in \mathring{U}_E(a)$, 因此, $f(x_n) \in V(A)$. 这样一来, 根据数列极限的定义推知 $\lim\limits_{n\to\infty} f(x_n) = A$.

现证命题的另一面. 如果 A 不是 $f(x)$ 当 $E \ni x \to a$ 时的极限, 则存在一个邻域 $V(A)$, 使得对于任何 $n \in \mathbb{N}$, 点 a 的 $\frac{1}{n}$ 邻域中必有点 $x_n \in E\backslash a$, 使得 $f(x_n) \notin V(A)$. 这正好是说, 虽然数列 $\{x_n\}$ 趋于 a, 数列 $\{f(x_n)\}$ 并不收敛于 A. ▶

2. 函数极限的性质 现在来建立一些经常用到的函数极限的性质, 其中大多数与已证过的数列极限的性质类似, 因此, 实际上我们已经都熟悉了. 此外, 根据刚才证明的命题 1, 函数极限的许多重要性质, 显然都能由数列极限的相应性质直接推出: 极限的唯一性, 极限的算术性质, 不等式中的极限过渡. 虽然如此, 我们还是给出了所有的证明. 即将看到, 这样做是有一定意义的.

我们希望读者注意, 要想建立函数极限的所有性质, 仅仅需要关于集合的极限点的去心邻域的两条性质:

B_1) $\mathring{U}_E(a) \neq \varnothing$, 即去心邻域是非空集;

B_2) $\forall \mathring{U}'_E(a) \forall \mathring{U}''_E(a) \exists \mathring{U}_E(a) (\mathring{U}_E(a) \subset \mathring{U}'_E(a) \cap \mathring{U}''_E(a))$, 即任意两个去心邻域之交还是一个去心邻域.

这个事实使我们形成了函数极限的一般概念. 并有可能在将来不只是对在数集上定义的函数, 而且对于一般函数都能使用极限理论. 为了使我们的叙述不是 §1 的简单重复, 我们在这里要用一些新的有益的方法和概念, 它们是 §1 中没有的.

a. 函数极限的一般性质 先给出几个定义.

定义 4 像前面一样, 把只取一个值的函数 $f: E \to \mathbb{R}$, 仍旧叫做*常值函数*. 设 a 是 E 的一个极限点, 如果在 a 的某个去心邻域 $\mathring{U}_E(a)$ 中, f 是常数, 就说函数 $f: E \to \mathbb{R}$ 当 $E \ni x \to a$ 时*最终为常值函数*.

定义 5 对于函数 $f: E \to \mathbb{R}$, 如果存在着数 $C \in \mathbb{R}$, 使得对于任何 $x \in E$ 成立

$$|f(x)| < C \text{或} f(x) < C \text{或} C < f(x),$$

①有时称这个命题为柯西极限定义 (邻域形式) 与海涅极限定义 (序列形式) 的等价性命题. 海涅 (H. E. Heine, 1821—1881) 是德国数学家.

就分别称 f 为有界函数, 上有界函数, 下有界函数.

如果以上三种关系中的某一个, 只在点 a 的某个去心邻域 $\mathring{U}_E(a)$ 中成立, 就分别把函数 $f: E \to \mathbb{R}$ 叫做当 $E \ni x \to a$ 时最终有界, 最终上有界, 最终下有界函数.

例 7 用公式 $f(x) = \sin\frac{1}{x} + x\cos\frac{1}{x}$ 定义的函数, 其中 $x \neq 0$, 在其定义域内不是有界函数, 但它是当 $x \to 0$ 时最终有界的函数.

例 8 上例之结论, 对于在 $\mathbb{R}$ 上定义的函数 $f(x) = x$ 也成立.

定理 1 a) ($f: E \to \mathbb{R}$ 当 $E \ni x \to a$时最终为常数 A) $\Rightarrow \left(\lim\limits_{E\ni x\to a} f(x) = A\right)$.

b) $\left(\exists \lim\limits_{E\ni x\to a} f(x)\right) \Rightarrow$ ($f: E \to \mathbb{R}$当 $E \ni x \to a$ 时最终有界).

c) $\left(\lim\limits_{E\ni x\to a} f(x) = A_1\right) \wedge \left(\lim\limits_{E\ni x\to a} f(x) = A_2\right) \Rightarrow (A_1 = A_2)$

◀ 最终为常数的函数有极限的结论a) 与有极限的函数必定最终有界的结论b), 由相应的定义直接推知. 现在我们来证明极限的唯一性.

假如 $A_1 \neq A_2$, 这时我们可以取得邻域 $V(A_1)$ 与 $V(A_2)$, 使它们没有公共点, 即 $V(A_1) \cap V(A_2) = \varnothing$. 根据极限定义, 有

$$\lim_{E\ni x\to a} f(x) = A_1 \Rightarrow \exists \mathring{U}'_E(a)(f(\mathring{U}'_E(a)) \subset V(A_1)),$$
$$\lim_{E\ni x\to a} f(x) = A_2 \Rightarrow \exists \mathring{U}''_E(a)(f(\mathring{U}''_E(a)) \subset V(A_2)).$$

今取 a 点 (E 的极限点) 的去心邻域 $\mathring{U}_E(a)$, 使 $\mathring{U}_E(a) \subset \mathring{U}'_E(a) \cap \mathring{U}''_E(a)$.(例如可以取 $\mathring{U}_E(a) = \mathring{U}'_E(a) \cap \mathring{U}''_E(a)$, 因为这个交也是去心邻域.)

因为 $\mathring{U}_E(a) \neq \varnothing$, 必能取得 $x \in \mathring{U}_E(a)$, 这时 $f(x) \in V(A_1) \cap V(A_2)$; 但这是不可能的, 因为, 由 $V(A_1), V(A_2)$ 的作法, 它们没有公共点. ▶

b. 极限过渡与算术运算

定义 6 如果两个数值函数 $f: E \to \mathbb{R}, g: E \to \mathbb{R}$ 有公共的定义域 E, 那么, 用以下各式

$$(f+g)(x) := f(x) + g(x),$$
$$(f \cdot g)(x) := f(x) \cdot g(x),$$
$$\left(\frac{f}{g}\right)(x) := \frac{f(x)}{g(x)}, \text{如果当} x \in E \text{时} g(x) \neq 0,$$

在 E 上定义的函数, 分别叫做它们的和、积与商.

定理 2 设 $f: E \to \mathbb{R}$ 与 $g: E \to \mathbb{R}$ 是有公共定义域的两个函数.

如果 $\lim\limits_{E\ni x\to a} f(x) = A$, $\lim\limits_{E\ni x\to a} g(x) = B$, 那么

a) $\lim\limits_{E\ni x\to a}(f+g)(x)=A+B$;

b) $\lim\limits_{E\ni x\to a}(f\cdot g)(x)=A\cdot B$;

c) $\lim\limits_{E\ni x\to a}\left(\dfrac{f}{g}\right)(x)=\dfrac{A}{B}$,假如 $B\neq 0$, 并且当 $x\in E$ 时, $g(x)\neq 0$.

这个定理, 正像第 2 段开始时所指出的, 如果应用第 1 段所证的命题, 就可以从数列极限的相应定理直接推出来.

也可以把数列极限的算术性质的证明过程重复一遍来证这个定理. 归结起来, 在证明中需要改变的是, 凡是以前写 "取 $N\in\mathbb{N}$, 从它开始 $\cdots$" 的地方, 现在就得改成选取 a 点在 E 中的某个去心邻域 $\mathring{U}(a)$. 建议读者来验证这一点.

而在这里, 我们将从这个定理当 $A=B=0$ 的简单特殊情形 (当然, 这时不考虑定理中的c)) 得出这个定理.

对于函数 $f:E\to\mathbb{R}$ 如果有 $\lim\limits_{E\ni x\to a}f(x)=0$, 就说 f 是当 $E\ni x\to a$ 时的无穷小.

命题 2 a) 设 $\alpha:E\to\mathbb{R},\beta:E\to\mathbb{R}$ 是当 $E\ni x\to a$ 时的无穷小, 那么, 其和$\alpha+\beta:E\to R$也是当$E\ni x\to 0$时的无穷小.

b) 设 $\alpha:E\to\mathbb{R},\beta:E\to\mathbb{R}$ 是当 $E\ni x\to a$ 时的无穷小, 那么, 其积 $\alpha\beta:E\to\mathbb{R}$ 也是当 $E\ni x\to a$ 时的无穷小.

c) 设 $\alpha:E\to\mathbb{R}$ 是当 $E\ni x\to a$ 时的无穷小, $\beta:E\ni x\to\mathbb{R}$ 当 $E\ni x\to a$ 时最终有界, 那么, 其积 $\alpha\cdot\beta:E\to\mathbb{R}$ 是当 $E\ni x\to a$ 时的无穷小

◀ a) 我们来验证

$$\left(\lim_{E\ni x\to a}\alpha(x)=0\right)\wedge\left(\lim_{E\ni x\to a}\beta(x)=0\right)$$
$$\Rightarrow\left(\lim_{E\ni x\to a}(\alpha+\beta)(x)=0\right).$$

设给定了 $\varepsilon>0$, 根据极限的定义知

$$\left(\lim_{E\ni x\to a}\alpha(x)=0\right)\Rightarrow\left(\exists\mathring{U}'_E(a)\forall x\in\mathring{U}'_E(a)\left(|\alpha(x)|<\frac{\varepsilon}{2}\right)\right),$$
$$\left(\lim_{E\ni x\to a}\beta(x)=0\right)\Rightarrow\left(\exists\mathring{U}''_E(a)\forall x\in\mathring{U}''_E(a)\left(|\beta(x)|<\frac{\varepsilon}{2}\right)\right).$$

因此, 对于去心邻域 $\mathring{U}_E(a)\subset\mathring{U}'_E(a)\cap\mathring{U}''_E(a)$, 得到

$$\forall x\in\mathring{U}_E(a)|(\alpha+\beta)(x)|=|\alpha(x)+\beta(x)|$$
$$\leqslant|\alpha(x)|+|\beta(x)|<\varepsilon.$$

这就验证了 $\lim\limits_{E\ni x\to a}(\alpha+\beta)(x)=0$.

b) 它是 c) 的特殊情形, 因为有极限的函数都是最终有界的.

c) 验证

$$\left(\lim_{E\in x\to a}\alpha(x)=0\right)\wedge(\exists M\in\mathbb{R}\exists\mathring{U}_E(a)\forall x\in\mathring{U}_E(a)(|\beta(x)|<M))$$
$$\Rightarrow\left(\lim_{E\ni x\to a}\alpha(x)\beta(x)=0\right).$$

设给定了 $\varepsilon>0$, 根据极限定义得

$$\left(\lim_{E\ni x\to a}\alpha(x)=0\right)\Rightarrow\left(\exists\mathring{U}'_E(a)\forall x\in\mathring{U}'_E(a)\left(|\alpha(x)|<\frac{\varepsilon}{M}\right)\right).$$

因此, 对于去心邻域 $\mathring{U}''_E(a)\subset\mathring{U}'_E(a)\cap\mathring{U}_E(a)$, 得到

$$\forall x\in\mathring{U}''_E(a)|(\alpha\cdot\beta)(x)|=|\alpha(x)\cdot\beta(x)|$$
$$=|\alpha(x)|\cdot|\beta(x)|<\frac{\varepsilon}{M}\cdot M=\varepsilon,$$

即验证了 $\lim\limits_{E\ni x\to a}(\alpha\cdot\beta)(x)=0$. ►

现在来作以下有益的注.

注

$$\boxed{\left(\lim_{E\ni x\to a}f(x)=A\right)\Leftrightarrow(f(x)=A+\alpha(x))\wedge\left(\lim_{E\ni x\to a}\alpha(x)=0\right).}$$

换句话说, 函数 $f:E\to\mathbb{R}$ 趋于 A, 当且仅当它能表成 $A+\alpha(x)$, 这里 $\alpha(x)$($f(x)$ 对 A 的偏差) 是当 $E\ni x\to a$ 时的无穷小函数①.

这可直接由极限的定义推出, 因为有 $\lim\limits_{E\ni x\to a}f(x)=A\Leftrightarrow\lim\limits_{E\ni x\to a}(f(x)-A)=0$.

现在我们根据这个注及无穷小函数的性质来证明关于函数极限算术性质的定理.

◄ a) 设 $\lim\limits_{E\ni x\to a}f(x)=A$ 且 $\lim\limits_{E\ni x\to a}g(x)=B$, 则 $f(x)=A+\alpha(x)$ 且 $g(x)=B+\beta(x)$, 其中 $\alpha(x),\beta(x)$ 是当 $E\ni x\to a$ 时的无穷小. 这时 $(f+g)(x)=f(x)+g(x)=A+\alpha(x)+B+\beta(x)=(A+B)+\gamma(x)$, 其中 $\gamma(x)=\alpha(x)+\beta(x)$, 作为无穷小之和, 是当 $E\ni x\to a$ 时的无穷小. 因此, $\lim\limits_{E\ni x\to a}(f+g)(x)=A+B$.

b) 仍把 $f(x),g(x)$ 表成 $f(x)=A+\alpha(x),g(x)=B+\beta(x)$ 的形式, 于是

$$(f\cdot g)(x)=f(x)\cdot g(x)=(A+\alpha(x))(B+\beta(x))=A\cdot B+\gamma(x),$$

其中 $\gamma(x)=A\cdot\beta(x)+B\cdot\alpha(x)+\alpha(x)\beta(x)$. 根据无穷小的性质, 它是当 $E\ni x\to a$ 时的无穷小函数.

① 花絮 这里的想法几乎是显然的, 但它在计算方面是很有益的, 在思想上是很重要的. 这种观念是法国数学家和力学家L. 卡诺(Carnot, L,1753—1823) 单独提出来的, 他是革命将领和院士, 是热力学创始人S 卡诺(1796—1832) 的父亲

因此, $\lim\limits_{E\ni x\to a}(f\cdot g)(x)=A\cdot B$.

c) 仍记 $f(x)=A+\alpha(x), g(x)=B+\beta(x)$, 其中 $\lim\limits_{E\ni x\to a}\alpha(x)=0$, $\lim\limits_{E\ni x\to a}\beta(x)=0$.

因为 $B\neq 0$, 存在去心邻域 $\mathring{U}(a)$, 使得在它的每一点上有 $|\beta(x)|<\dfrac{|B|}{2}$, 从而 $g(x)=|B+\beta(x)|\geqslant B-|\beta(x)|>\dfrac{|B|}{2}$, 这时, 在 $\mathring{U}_E(a)$ 中也有 $\dfrac{1}{|g(x)|}<\dfrac{2}{|B|}$, 亦即, 函数 $\dfrac{1}{|g(x)|}$ 当 $E\ni x\to a$ 时最终有界, 现在记

$$\left(\frac{f}{g}\right)(x)-\frac{A}{B}=\frac{f(x)}{g(x)}-\frac{A}{B}=\frac{A+\alpha(x)}{B+\beta(x)}-\frac{A}{B}$$
$$=\frac{1}{g(x)}\cdot\frac{1}{B}(B\cdot\alpha(x)-A\cdot\beta(x))=\gamma(x).$$

根据无穷小的性质 (以及刚证的 $\dfrac{1}{g(x)}$ 的最终有界性), 可知函数 $\gamma(x)$ 是当 $E\ni x\to a$ 时的无穷小, 于是证明了 $\lim\limits_{E\ni x\to a}\left(\dfrac{f}{g}\right)(x)=\dfrac{A}{B}$ ►

c. 极限过渡与不等式

定理 3 a) *设函数 $f:E\to\mathbb{R}, g:E\to\mathbb{R}$ 满足 $\lim\limits_{E\ni x\to a}f(x)=A$ $\lim\limits_{E\ni x\to a}g(x)=B$ 且 $A<B$, 那么, 必存在点 a 在集合 E 中的一个去心邻域 $\mathring{U}_E(a)$, 在其中的每个点上, 不等式 $f(x)<g(x)$ 成立.*

b) 如果在集合 E 上定义的三个函数 $f:E\to\mathbb{R}, g:E\to\mathbb{R}, h:E\to\mathbb{R}$ 之间, 不等式 $f(x)\leqslant g(x)\leqslant h(x)$ 成立, 并且 $\lim\limits_{E\ni x\to a}f(x)=\lim\limits_{E\ni x\to a}h(x)=C$, 那么, 当 $E\ni x\to a$ 时, $g(x)$ 也有极限, 并且

$$\lim_{E\ni x\to a}g(x)=C.$$

◄ a) 取数 C 使 $A<C<B$. 据极限之定义, 必有点 a 在集合 E 中的去心邻域 $\mathring{U}'_E(a)$ 及 $\mathring{U}''_E(a)$, 使得当 $x\in\mathring{U}'_E(a)$ 时, $|f(x)-A|<C-A$, 而当 $x\in\mathring{U}''_E(a)$ 时, $|g(x)-B|<B-C$. 因此, 在含于 $\mathring{U}'_E(a)\cap\mathring{U}''_E(a)$ 的任何去心邻域 $\mathring{U}_E(a)$ 中, 必有

$$f(x)<A+(C-A)=C=B-(B-C)<g(x).$$

b) 设 $\lim\limits_{E\ni x\to a}f(x)=\lim\limits_{E\ni x\to a}h(x)=C$, 那么, 对任何固定的 $\varepsilon>0$, 可求得点 a 在集合 E 中的去心邻域 $\mathring{U}'_E(a)$ 与 $\mathring{U}''_E(a)$, 使得当 $x\in\mathring{U}'_E(a)$ 时, $C-\varepsilon<f(x)$, 当 $x\in\mathring{U}''_E(a)$ 时, $h(x)<C+\varepsilon$. 因此, 在包含在 $\mathring{U}'_E(a)\cap\mathring{U}''_E(a)$ 中的任何去心邻域 $\mathring{U}_E(a)$ 中, 必有 $C-\varepsilon<f(x)\leqslant g(x)\leqslant h(x)<C+\varepsilon$, 即 $|g(x)-C|<\varepsilon$, 因之 $\lim\limits_{E\ni x\to a}g(x)=C$. ►

推论　设 $\lim_{E\ni x\to a} f(x) = A$, $\lim_{E\ni x\to a} g(x) = B$. 如果在点 a 的某个去心邻域 $\mathring{U}_E(a)$ 中

a) 满足 $f(x) > g(x)$, 那么 $A \geqslant B$;

b) 满足 $f(x) \geqslant g(x)$, 那么, $A \geqslant B$;

c) 满足 $f(x) > B$, 那么 $A \geqslant B$;

d) 满足 $f(x) \geqslant B$, 那么 $A \geqslant B$.

◀ 由定理之结论a), 用反证法直接推出结论a) 与b) 来; 结论c) 与d) 是a),b) 当 $g(x) \equiv B$ 时的特例. ▶

d. 两个重要极限　　在进一步阐述函数极限的理论之前, 我们先用两个重要例子来展示已经证明了的定理的应用.

例 9

$$\boxed{\lim_{x\to 0}\frac{\sin x}{x} = 1.}$$

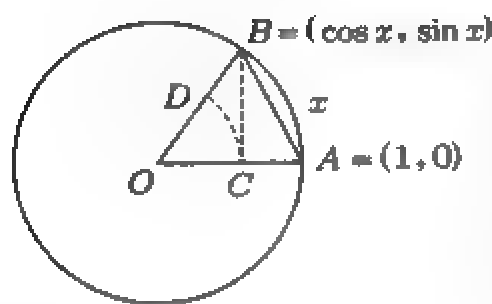


图 8

这里, 我们将借助于中学里对 $\sin x$ 所下的定义, 即把它定义作点 (1,0)(以坐标原点为中心) 旋转 x(弧度) 角所得点的纵标. 这个定义的圆满程度, 完全依赖于在多大程度上建立了旋转与实数之间的联系. 因为实数系本身在中学里没有得到足够细致的描述, 必须承认, 有必要把$\sin x$的定义精确化 (对$\cos x$也是如此).

我们将在适当的时候做这件事, 并论证以下断言. 而现在这些断言的论证将依据直观进行.

a) 证明

$$\cos^2 x < \frac{\sin x}{x} < 1,\ 这里\ 0 < |x| < \frac{\pi}{2}.$$

◀ 因为 $\cos^2 x$ 与 $\frac{\sin x}{x}$ 都是偶函数, 所以只需讨论 $0 < x < \frac{\pi}{2}$ 的情形. 由图 8 与 $\cos x$ 及 $\sin x$ 的定义, 比较扇形 $\sphericalangle OCD$, 三角形 ΔOAB 与扇形 $\sphericalangle OAB$ 的面积, 得到

$$\begin{aligned}
S_{\sphericalangle OCD} &= \frac{1}{2}|OC|\cdot|\overset{\frown}{CD}| = \frac{1}{2}(\cos x)(x\cos x) = \frac{1}{2}x\cos^2 x < \\
S_{\Delta OAB} &= \frac{1}{2}|OA|\cdot|BC| = \frac{1}{2}(1)(\sin x) = \frac{1}{2}\sin x < S_{\sphericalangle OAB} \\
&= \frac{1}{2}|OA|\cdot|\overset{\frown}{AB}| = \frac{1}{2}\cdot 1\cdot x = \frac{1}{2}x.
\end{aligned}$$

将此不等式除以 $\frac{1}{2}x$, 即得所证之结论. ▶

b) 由 a) 推出, 对于任何 $x \in \mathbb{R}$,

$$|\sin x| \leqslant |x|,$$

并且只有当 $x=0$ 时等式才成立.

◀ 当 $0<|x|<\pi/2$ 时,a) 中已证

$$|\sin x| < |x|.$$

但 $|\sin x| \leqslant 1$, 所以当 $|x| \geqslant \dfrac{\pi}{2} > 1$ 时, 上面不等式也成立. 只有当 $x=0$ 时, 才有 $\sin x = x = 0$. ▶

c) 由 b) 推出

$$\lim_{x\to 0} \sin x = 0.$$

◀ 因为 $0 \leqslant |\sin x| \leqslant |x|$, 所以根据函数极限与不等式关系的定理推知 $\lim\limits_{x\to 0} |\sin x| = 0$, 因此, $\lim\limits_{x\to 0} \sin x = 0$. ▶

d) 现在证明 $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$.

◀ 不妨认为 $|x| < \dfrac{\pi}{2}$. 根据 a) 中所得不等式, 得到

$$1-\sin^2 x < \frac{\sin x}{x} < 1.$$

但

$$\lim_{x\to 0}(1-\sin^2 x) = 1-(\lim_{x\to 0}\sin x)(\lim_{x\to 0}\sin x) = 1-0 = 1,$$

因此, 根据关于不等式中极限过渡的定理, 即可推知 $\lim\limits_{x\to 0} \dfrac{\sin x}{x} = 1$. ▶

例 10 *根据极限理论定义指数函数、对数函数与幂函数.* 我们现在就来展示, 在已经有了实数理论和极限理论的条件下, 对中学数学中的指数函数与对数函数的定义能作哪些补充和怎样补充.

为了便于今后的引证和整个理论画面的完整性, 我们拟从头做起.

a) *指数函数.* 设 $a>1$.

1° 对于 $n \in \mathbb{N}$, 用归纳法令

$$a^1 := a, a^{n+1} := a^n \cdot a.$$

这样, 就能定义在 $\mathbb{N}$ 上的函数 a^n, 由此定义可见, 这个函数具有性质: 当 $m, n \in \mathbb{N}$ 且 $m>n$ 时,

$$\frac{a^m}{a^n} = a^{m-n}.$$

2° 根据这个性质, 自然地引导出定义

$$a^0 := 1, \qquad a^{-n} := \frac{1}{a^n} \text{ 当 } n \in \mathbb{N},$$

这样做了之后, 函数 a^n 的定义就被推广到整数集 $\mathbb{Z}$ 上了, 并且当 $m, n \in \mathbb{Z}$ 时, 有

$$a^m \cdot a^n = a^{m+n}.$$

3° 我们在实数理论中已经看到, 当 $a > 0$ 且 $n \in \mathbb{N}$ 时, 存在着 a 的唯一的 n 次算术根, 即存在一数 $x > 0$, 使得 $x^n = a$. 这个 x 记成 $x = a^{\frac{1}{n}}$. 如果我们希望保留指数的加法规则

$$a = a^1 = (a^{\frac{1}{n}})^n = a^{\frac{1}{n}} \cdot \cdots \cdot a^{\frac{1}{n}} = a^{\frac{1}{n}+\cdots+\frac{1}{n}}$$

的话, 这种记法是合适的.

根据同样的原因, 对于 $n \in \mathbb{N}$ 及 $m \in \mathbb{Z}$, 令

$$a^{\frac{m}{n}} := (a^{\frac{1}{n}})^m \text{ 和 } a^{-\frac{1}{n}} := (a^{\frac{1}{n}})^{-1}$$

是自然的. 如果对 $k \in \mathbb{Z}$ 有 $a^{\frac{mk}{nk}} = a^{\frac{m}{n}}$, 就可认为已对 $r \in \mathbb{Q}$ 定义了 a^r.

4° 当 $0 < x, 0 < y$ 时, 用归纳法可验证, 当 $n \in \mathbb{N}$ 时,

$$(x < y) \Leftrightarrow x^n < y^n,$$

因此, 可以推知

$$(x = y) \Leftrightarrow x^n = y^n.$$

5° 现在可以证明有理指数的运算规则, 其中包括

$$a^{\frac{mk}{nk}} = a^{\frac{m}{n}} \quad \text{当} k \in \mathbb{Z} \text{时},$$

及

$$a^{\frac{m_1}{n_1}} \cdot a^{\frac{m_2}{n_2}} = a^{\frac{m_1}{n_1}+\frac{m_2}{n_2}}.$$

◀ 实际上, $a^{\frac{mk}{nk}} > 0$ 且 $a^{\frac{m}{n}} > 0$. 其次, 由于

$$(a^{\frac{mk}{nk}})^{nk} = ((a^{\frac{1}{nk}})^{mk})^{nk} = (a^{\frac{1}{nk}})^{mk\,nk}$$
$$= ((a^{\frac{1}{nk}})^{nk})^{mk} = a^{mk}$$

及

$$(a^{\frac{m}{n}})^{nk} = ((a^{\frac{1}{n}})^n)^{mk} = a^{mk},$$

所以, 根据 4°, 得到了要证的第一个等式.

类似地

$$\begin{aligned}(a^{\frac{m_1}{n_1}} \cdot a^{\frac{m_2}{n_2}})^{n_1n_2} &= (a^{\frac{m_1}{n_1}})^{n_1n_2} \cdot (a^{\frac{m_2}{n_2}})^{n_1n_2} \\ &= ((a^{\frac{1}{n_1}})^{n_1})^{m_1n_2} \cdot ((a^{\frac{1}{n_2}})^{n_2})^{m_2n_1} = a^{m_1n_2} \cdot a^{m_2n_1} \\ &= a^{m_1n_2+m_2n_1},\end{aligned}$$

及

$$\begin{aligned}(a^{\frac{m_1}{n_1}+\frac{m_2}{n_2}})^{n_1n_2} &= (a^{\frac{m_1n_2+m_2n_1}{n_1n_2}})^{n_1n_2} = ((a^{\frac{1}{n_1n_2}})^{n_1n_2})^{m_1n_2+m_2n_1} \\ &= a^{m_1n_2+m_2n_1},\end{aligned}$$

所以第二个等式也证明了. ►

这样, 我们对 $r \in \mathbb{Q}$ 定义了 a^r 而且, 对于任何 $r \in \mathbb{Q}$ 有 $a^r > 0$ 及对于任何 $r_1, r_2 \in \mathbb{Q}$ 有

$$a^{r_1} \cdot a^{r_2} = a^{r_1+r_2}.$$

6° 由 4° 推出, 当 $r_1, r_2 \in \mathbb{Q}$ 时,

$$(r_1 < r_2) \Rightarrow (a^{r_1} < a^{r_2}).$$

◄ 因为 $(1 < a) \Leftrightarrow (1 < a^{\frac{1}{n}})$ 对一切 $n \in \mathbb{N}$ 成立, 这由 4° 立刻可以推出来, 因此, 仍由 4° 推知, 当 $n, m \in \mathbb{N}$ 时, $(a^{\frac{1}{n}})^m = a^{\frac{m}{n}} > 1$. 所以, 当 $1 < a$ 时, 对于使 $r > 0$ 的 $r \in \mathbb{Q}$ 得到 $a^r > 1$.

而当 $r_1 < r_2$ 时, 据 5° 就得到

$$a^{r_2} = a^{r_1} \cdot a^{r_2 - r_1} > a^{r_1} \cdot 1 = a^{r_1}. ►$$

7° 设 $r_0 \in \mathbb{Q}$, 则有

$$\lim_{\mathbb{Q} \ni r \to r_0} a^r = a^{r_0}.$$

◄ 可以验证, 当 $\mathbb{Q} \ni p \to 0$ 时, $a^p \to 1$. 因为, 当 $|p| < \dfrac{1}{n}$ 时, 由 6° 能得到

$$a^{-\frac{1}{n}} < a^p < a^{\frac{1}{n}}.$$

我们知道, 当 $n \to \infty$ 时, $a^{\frac{1}{n}} \to 1$(及 $a^{-\frac{1}{n}} \to 1$) . 这时, 用标准的证法即可验证, 对 $\varepsilon > 0$, 存在 $\delta > 0$, 使得当 $|p| \leqslant \delta$ 时, 有

$$1 - \varepsilon < a^p < 1 + \varepsilon.$$

可取 $\dfrac{1}{n}$ 作为 δ, 使 $1 - \varepsilon < a^{-\frac{1}{n}}$ 且 $a^{\frac{1}{n}} < 1 + \varepsilon$.

现在来证明基本命题.

对于 $\varepsilon>0$, 选择 δ 使得当 $|p|<\delta$ 时

$$1-\varepsilon a^{-r_0}<a^p<1+\varepsilon a^{-r_0}.$$

今若 $|r-r_0|<\delta$, 则

$$a^{r_0}(1-\varepsilon a^{-r_0})<a^r=a^{r_0}\cdot a^{r-r_0}<a^{r_0}(1+\varepsilon a^{-r_0}),$$

或

$$a^{r_0}-\varepsilon<a^r<a^{r_0}+\varepsilon. \blacktriangleright$$

这样, 在 $\mathbb{Q}$ 上定义了具有以下性质的函数 a^r:

$$\begin{aligned}a^1=a>1;\\ a^{r_1}\cdot a^{r_2}=a^{r_1+r_2};\\ \text{当}r_1<r_2\text{时}, a^{r_1}<a^{r_2};\\ \text{当}\mathbb{Q}\ni r_1\to r_2\text{时}, a^{r_1}\to a^{r_2}.\end{aligned}$$

现在我们用以下方法, 把这个函数开拓到整个数轴上去.

8° 设 $x\in\mathbb{R}$, $s=\sup\limits_{\mathbb{Q}\ni r<x}a^r, i=\inf\limits_{\mathbb{Q}\ni r>x}a^r$, 显然 $s,i\in\mathbb{R}$, 因为当 $r_1<x<r_2$ 时, 有 $a^{r_1}<a^{r_2}$.

我们要证明, 实际上有 $s=i$(这时我们就用 a^x 表示这个数).

◀ 据 s 与 i 之定义, 当 $r_1<x<r_2$ 时, 有

$$a^{r_1}\leqslant s\leqslant i\leqslant a^{r_2}.$$

因此, $0\leqslant i-s\leqslant a^{r_2}-a^{r_1}=a^{r_1}(a^{r_2-r_1}-1)<s(a^{r_2-r_1}-1)$. 但当 $\mathbb{Q}\ni p\to 0$ 时 $a^p\to 1$, 所以对于任何 $\varepsilon>0$, 能求得 $\delta>0$, 使得当 $0<r_2-r_1<\delta$ 时, 有 $a^{r_2-r_1}-1<\varepsilon/s$. 这时, 我们得到 $0\leqslant i-s<\varepsilon$, 再由 $\varepsilon>0$ 的任意性, 得知 $i=s$. ▶

令　$a^x:=s=i$.

9° 今证 $a^x=\lim\limits_{\mathbb{Q}\ni r\to x}a^r$.

◀ 从 8° 看到, 对 $\varepsilon>0$, 存在 $r'<x$, 使得 $s-\varepsilon<a^{r'}\leqslant s=a^x$, 又存在 r'' 使 $a^x=i\leqslant a^{r''}<i+\varepsilon$, 因为, 当 $r'<r<r''$ 时, 必有 $a^{r'}<a^r<a^{r''}$, 所以, 对于开区间 $]r',r''[$ 中的 $r\in\mathbb{Q}$, 也必定有

$$a^x-\varepsilon<a^r<a^x+\varepsilon. \blacktriangleright$$

现在来讨论这样构造出来的 $\mathbb{R}$ 上的函数 a^x 的性质.

10° 设 $x_1,x_2\in\mathbb{R}$, 当 $a>1$ 时, $(x_1<x_2)\Rightarrow(a^{x_1}<a^{x_2})$.

◀ 在开区间 $]x_1, x_2[$ 上, 存在两个有理数 $r_1 < r_2$. 既然 $x_1 < r_1 < r_2 < x_2$, 那么, 据 8° 中给出的 a^x 的定义及函数 a^r 在 $\mathbb{Q}$ 上的性质, 就得到

$$a^{x_1} < a^{r_1} < a^{r_2} < a^{x_2}. ▶$$

11° 对任何 $x_1, x_2 \in \mathbb{R}$, $a^{x_1} \cdot a^{x_2} = a^{x_1+x_2}$ 成立.

◀ 根据乘积的绝对误差的估值及性质 9°, 可以断定, 对于任何 $\varepsilon > 0$, 必能找到 $\delta' > 0$, 使得当 $|x_1 - r_1| < \delta', |x_2 - r_2| < \delta'$ 时,

$$a^{x_1} \cdot a^{x_2} - \frac{\varepsilon}{2} < a^{r_1} \cdot a^{r_2} < a^{x_1} \cdot a^{x_2} + \frac{\varepsilon}{2}.$$

如果需要, 把 δ' 再缩小到 $\delta, \delta < \delta'$, 使得当 $|x_1 - r_1| < \delta, |x_2 - r_2| < \delta$, 从而 $|(x_1 + x_2) - (r_1 + r_2)| < 2\delta$ 时,

$$a^{r_1+r_2} - \frac{\varepsilon}{2} < a^{x_1+x_2} < a^{r_1+r_2} + \frac{\varepsilon}{2}.$$

但是, 当 $r_1, r_2 \in \mathbb{Q}$ 时, $a^{r_1} \cdot a^{r_2} = a^{r_1+r_2}$, 因此

$$a^{x_1} \cdot a^{x_2} - \varepsilon < a^{x_1+x_2} < a^{x_1} \cdot a^{x_2} + \varepsilon.$$

由 ε 的任意性, 推知

$$a^{x_1} \cdot a^{x_2} = a^{x_1+x_2}. ▶$$

12° $\lim\limits_{x \to x_0} a^x = a^{x_0}$(注意,"$x \to x_0$" 是 "$\mathbb{R} \ni x \to x_0$" 的简记法)

◀ 首先验证 $\lim\limits_{x \to 0} a^x = 1$. 对 $\varepsilon > 0$ 求得 $n \in \mathbb{N}$, 使得

$$1 - \varepsilon < a^{-\frac{1}{n}} < a^{\frac{1}{n}} < 1 + \varepsilon.$$

这时, 由 10°, 当 $|x| < \dfrac{1}{n}$ 时将有

$$1 - \varepsilon < a^{-\frac{1}{n}} < a^x < a^{\frac{1}{n}} < 1 + \varepsilon,$$

即验证了 $\lim\limits_{x \to 0} a^x = 1$.

现在取 $\delta > 0$, 使得当 $|x - x_0| < \delta$ 时, $|a^{x-x_0} - 1| < \varepsilon a^{-x_0}$, 就得到

$$a^{x_0} - \varepsilon < a^x = a^{x_0}(a^{x-x_0} - 1) < a^{x_0} + \varepsilon.$$

即证明了 $\lim\limits_{x \to x_0} a^x = a^{x_0}$. ▶

13° 证明函数 $x \mapsto a^x$ 的值集是一切正实数之集 $\mathbb{R}_+$.

◀ 令 $y_0 \in \mathbb{R}_+$. 如果 $a > 1$, 我们知道, 存在 $n \in \mathbb{N}$, 使得 $a^{-n} < y_0 < a^n$. 因此,

$$A = \{x \in \mathbb{R} | a^x < y_0\} \quad 与 \quad B = \{x \in \mathbb{R} | y_0 < a^x\}$$

这两个集合都不空. 但是, 由于 (当 $a>1$ 时)$(x_1<x_2)\Leftrightarrow a^{x_1}<a^{x_2}$, 那么, 对于 $x_1\in A,x_2\in B,x_1,x_2\in\mathbb{R}$, 就必定有 $x_1<x_2$. 因此, 对于集合 A,B 应用完备公理, 就得到一数 x_0, 使对于任何 $x_1\in A$ 及 $x_2\in B$, 有 $x_1\leqslant x_0\leqslant x_2$. 今证 $a^{x_0}=y_0$.

假如 $a^{x_0}<y_0$, 那么, 由于当 $n\to\infty$ 时, $a^{x_0+\frac{1}{n}}\to a^{x_0}$, 应有数 $n\in\mathbb{N}$ 使 $a^{x_0+\frac{1}{n}}<y_0$. 这就得到 $\left(x_0+\dfrac{1}{n}\right)\in A$. 但这与已经证明的 x_0 分割 A 与 B 的性质不相容. 所以 $a^{x_0}<y_0$ 之假定不能成立. 同样可证不等式 $y_0<a^{x_0}$ 也不可能成立. 因此根据实数的性质, 必然 $a^{x_0}=y_0$. ▶

14° 我们一直都假定了 $a>1$, 但是, 以上所作的一切论述, 对于 $0<a<1$ 的情形可照样进行. 在这种条件下, 如果 $r>0$, 则有 $0<a^r<1$. 由此, 在 6° 中, 以及最后在 10° 中, 现在得到的应是: 当 $0<a<1$ 时, $(x_1<x_2)\Rightarrow(a^{x_1}>a^{x_2})$.

因此, 对 $a>0,a\neq1$, 我们建立了在实数集上定义的指数函数$x\mapsto a^x$, 它具有以下的性质:

1) $a^1=a$;

2) $a^{x_1}\cdot a^{x_2}=a^{x_1+x_2}$;

3) 当 $x\to x_0$ 时, $a^x\to a^{x_0}$;

4) 若 $a>1$, 则 $(a^{x_1}<a^{x_2})\Leftrightarrow(x_1<x_2)$,
若 $0<a<1$, 则 $(a^{x_1}>a^{x_2})\Leftrightarrow(x_1<x_2)$;

5) 函数 $x\mapsto a^x$ 的值集是一切正实数之集

$$\mathbb{R}_+=\{y\in\mathbb{R}|0<y\}.$$

定义 7　映射 $x\mapsto a^x$ 叫做以 a 为底的*指数函数*. 当 $a=e$ 时, 这个经常遇到的函数 $x\mapsto e^x$ 常记作 $\exp x$. 与此相关, 函数 $x\mapsto a^x$ 有时也记作 $\exp_a x$.

b) *对数函数*. 由指数函数的性质知映射 $\exp_a:\mathbb{R}\to\mathbb{R}_+$ 是双射, 所以它有反函数.

定义 8　$\exp_a:\mathbb{R}\to\mathbb{R}_+$ 的逆映射叫做*以 $a(0<a,a\neq1)$ 为底的对数函数*, 记作

$$\log_a:\mathbb{R}_+\to\mathbb{R}.$$

定义 9　以 $a=e$ 为底的对数函数, 或对数, 叫做*自然对数*, 并记作 $\ln:\mathbb{R}_+\to\mathbb{R}$.

可用另一种方法引入对数, 从它能清楚地看出对数这个述语的来由. 这另一种方法, 甚至更自然、更明显, 我们将在建立了微积分初步之后, 再来叙述它.

根据定义, 对数函数是指数函数的反函数, 于是

$$\forall x\in\mathbb{R}(\log_a(a^x)=x),$$
$$\forall y\in\mathbb{R}_+(a^{\log_a y}=y).$$

由这个定义及指数函数的性质得到, 对数在自己的定义域 $\mathbb{R}_+$ 中有如下性质:

1′) $\log_a a = 1$;

2′) $\log_a(y_1 \cdot y_2) = \log_a y_1 + \log_a y_2$;

3′) 当 $\mathbb{R}_+ \ni y \to y_0 \in \mathbb{R}_+$ 时, $\log_a y \to \log_a y_0$;

4′) $a > 1$ 时, $(\log_a y_1 < \log_a y_2) \Leftrightarrow (y_1 < y_2)$;

$0 < a < 1$ 时, $(\log_a y_1 > \log_a y_2) \Leftrightarrow (y_1 < y_2)$;

5′) 函数 $\log_a : \mathbb{R}_+ \to \mathbb{R}$ 的值集为整个实数集 $\mathbb{R}$.

◀ 由指数函数之性质 1) 与对数之定义即得 1′).

由指数函数之性质 2) 得到 2′) 实际上, 设 $x_1 = \log_a y_1, x_2 = \log_a y_2$. 于是 $y_1 = a^{x_1}, y_2 = a^{x_2}$, 而且, 由 2), 得 $y_1 \cdot y_2 = a^{x_1} \cdot a^{x_2} = a^{x_1+x_2}$, 因此, $\log_a(y_1 \cdot y_2) = x_1 + x_2$.

类似地, 从指数函数的性质 4), 可导出对数函数的性质 4′).

显然 5) $\Rightarrow$ 5′).

还有 3′) 待证明.

据对数的性质 2′)

$$\log_a y - \log_a y_0 = \log_a \left(\frac{y}{y_0}\right),$$

所以不等式

$$-\varepsilon < \log_a y - \log_a y_0 < \varepsilon$$

与

$$\log_a(a^{-\varepsilon}) = -\varepsilon < \log_a \left(\frac{y}{y_0}\right) < \varepsilon = \log_a(a^{\varepsilon})$$

等价. 据对数的性质 4′), 后者又与

$$a^{-\varepsilon} < \frac{y}{y_0} < a^{\varepsilon} \qquad (\text{当} a > 1 \text{时})$$
$$a^{\varepsilon} < \frac{y}{y_0} < a^{-\varepsilon} \qquad (\text{当} 0 < a < 1 \text{时})$$

等价.

不论是哪种情形, 亦即不论

$$y_0 a^{-\varepsilon} < y < y_0 a^{\varepsilon} \qquad (\text{对} a > 1)$$

还是

$$y_0 a^{\varepsilon} < y < y_0 a^{-\varepsilon} \qquad (\text{对} 0 < a < 1)$$

都得到

$$-\varepsilon < \log_a y - \log_a y_0 < \varepsilon.$$

这就证明了

$$\lim_{\mathbb{R}_+ \ni y \to y_0 \in \mathbb{R}_+} \log_a y = \log_a y_0. \blacktriangleright$$

在图 9 中画出了函数 $e^x, 10^x, \ln x, \log_{10} x =: \log x$ 等的图像, 而在图 10 中画出了 $\left(\frac{1}{e}\right)^x, (0.1)^x, \log_{e^{-1}} x, \log_{0.1} x$ 的图像.

还有一个经常用到的对数性质, 即

对任何 $b > 0$ 及任何 $\alpha \in \mathbb{R}$, 以下等式成立:

6′) $\log_a(b^\alpha) = \alpha \log_a b$.

◀1° 当 $\alpha = n \in \mathbb{N}$ 时, 等式成立, 因为由对数性质 2′), 用归纳法可得

$$\log_a(y_1 \cdot \cdots \cdot y_n) = \log_a y_1 + \cdots + \log_a y_n,$$

于是

$$\log_a(b^n) = \log_a b + \cdots + \log_a b = n \log_a b.$$

2° $\log_a(b)^{-1} = -\log_a b$, 因为, 如果令 $\beta = \log_a b$, 那么

$$b = a^\beta, b^{-1} = a^{-\beta} \text{且} \log_a(b)^{-1} = -\beta.$$

3° 由 1° 和 2° 就知道对于 $\alpha \in \mathbb{Z}$ 等式 $\log_a(b^\alpha) = \alpha \log_a b$ 成立.

4° 当 $n \in \mathbb{Z}$ 时, $\log_a(b^{\frac{1}{n}}) = \frac{1}{n} \log_a b$.

实际上

$$\log_a b = \log_a(b^{\frac{1}{n}})^n = n \log_a(b^{\frac{1}{n}}).$$

5° 对任何有理数 $\alpha = \frac{m}{n} \in \mathbb{Q}$, 现在可验证命题正确. 实际上,

$$\frac{m}{n} \log_a b = m \log_a b^{\frac{1}{n}} = \log_a(b^{\frac{1}{n}})^m = \log_a(b^{\frac{m}{n}}).$$

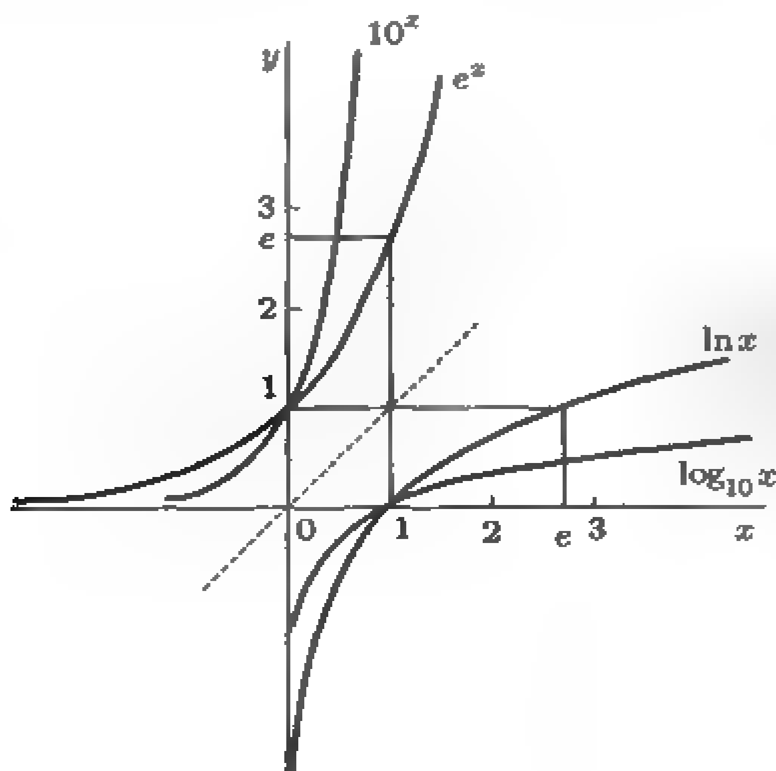


图 9

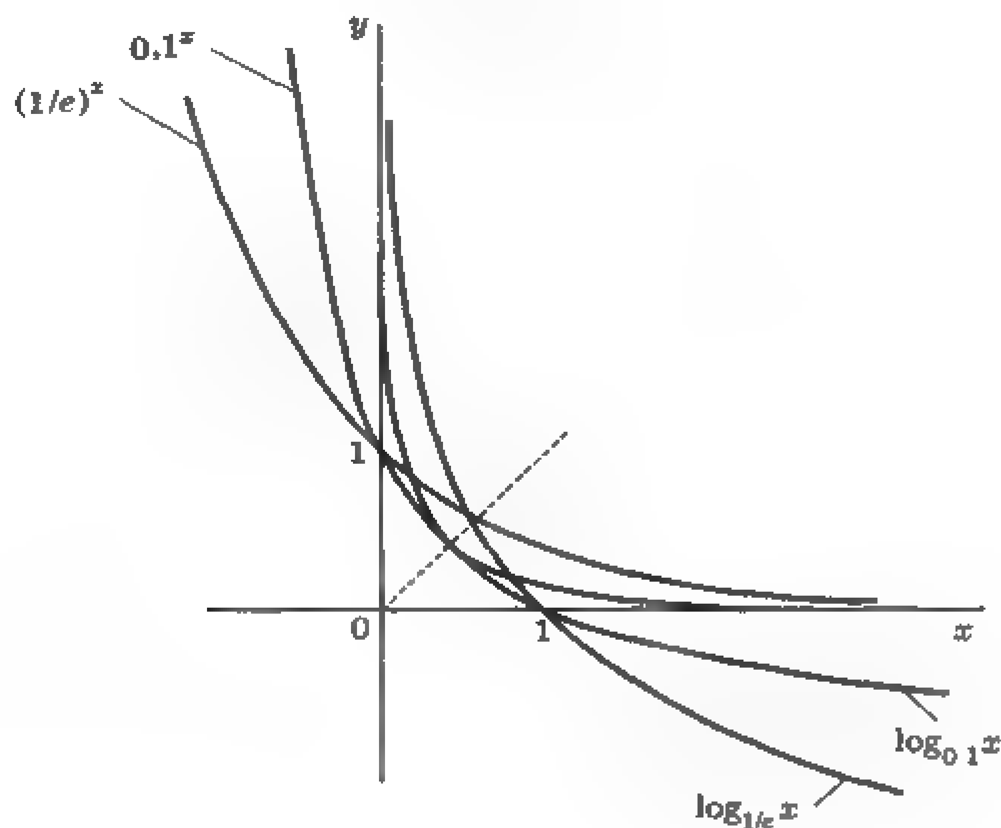


图 10

6° 既已对于任何有理数 $r \in \mathbb{Q}$ 证明了等式

$$\log_a b^r = r \log_a b,$$

令 r 沿着 $\mathbb{Q}$ 趋近于 α, 据指数函数的性质 3), 与对数函数的性质 3′), 就得知当 r 足够接近于 α 时, b^r 就接近于 b^α, 于是 $\log_a b^r$ 就接近于 $\log_a b^\alpha$. 这就是

$$\lim_{\mathbb{Q} \ni r \to \alpha} \log_a b^r = \log_a b^\alpha.$$

但 $\log_a b^r = r \log_a b$, 所以

$$\log_a b^\alpha = \lim_{\mathbb{Q} \ni r \to \alpha} \log_a b^r = \lim_{\mathbb{Q} \ni r \to \alpha} r \log_a b = \alpha \log_a b. \blacktriangleright$$

我们约定 $\forall \alpha \in \mathbb{R} (1^\alpha = 1)$.

由已证明的这些对数性质, 又可得到以下的结论: 对于任何 $\alpha, \beta \in \mathbb{R}$ 及 $\alpha > 0$, 下面等式成立:

6) $(a^\alpha)^\beta = a^{\alpha\beta}$.

◀ 当 $a = 1$ 时, 由于对任何 α 都有 $1^\alpha = 1$, 等式显然成立.

当 $a \neq 1$ 时, 则有

$\log_a((a^\alpha)^\beta) = \beta \log_a(a^\alpha) = \beta \cdot \alpha \log_a a = \alpha \cdot \beta = \log_a(a^{\alpha\beta})$, 再由对数性质 4′), 就得 $(a^\alpha)^\beta = a^{\alpha\beta}$. ▶

c) 幂函数. 我们已对任何 $x > 0$ 及任何 $\alpha \in \mathbb{R}$ 定义了 x^α(读做 x 的 α 次幂).

定义 10 定义在正数集 $\mathbb{R}_+$ 上的函数 $x \mapsto x^\alpha$, 叫做幂函数, 称 α 为幂指数.

幂函数显然是指数函数与对数函数的复合, 确切地说就是

$$x^{\alpha} = a^{\log_a(x^{\alpha})} = a^{\alpha \log_a x}$$

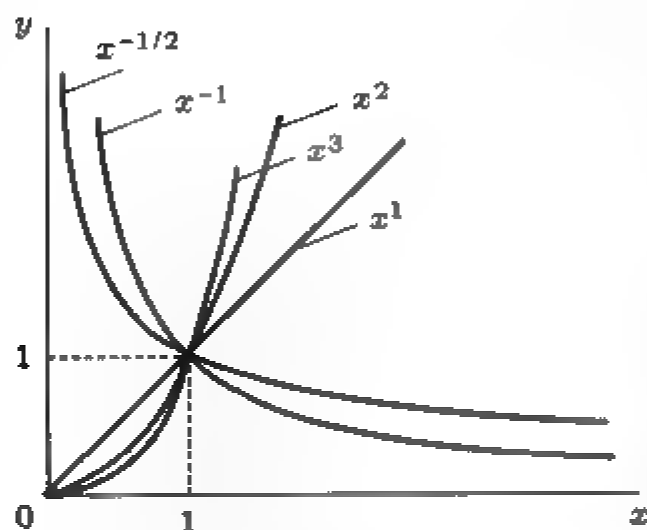


图 11

在图 11 中, 对一些幂指数画出了 $y = x^{\alpha}$ 的图像.

3. 函数极限的一般定义 (对基的极限)

在第二段中, 对于定义在去心邻域中的函数, 我们证明了一些极限性质. 不难看出, 在这些证明中, 除了那里引进的去心邻域的性质 B_1) 和 B_2), 实际上不再需要别的什么, 因此, 有理由引进以下数学对象.

a. 基的定义与主要例子

定义 11 由集合 X 的某些子集 $B \subset X$ 组成的集族 $\mathcal{B}$ 称为集合 X 中的基, 假如它满足两个条件:

B_1) $\forall B \in \mathcal{B} (B \neq \varnothing)$;

B_2) $\forall B_1 \in \mathcal{B} \forall B_2 \in \mathcal{B} \exists B \in \mathcal{B} (B \subset B_1 \cap B_2)$.

换句话说, 族 $\mathcal{B}$ 的元素不是空集, 并且任二元素之交都含有族的某个元素.

现在把分析中最常用的基列表如下:

基的表示法	表示法的名称	由什么集合 (元素) 组成的基	基中的元素的定义与表示法
$x \to a$	x 趋于 a	点 $a \in \mathbb{R}$ 的去心邻域组成的基	$\mathring{U}(a) := \{x \in \mathbb{R} \mid a - \delta_1 < x < a + \delta_2 \wedge x \neq a\}$ 其中 $\delta_1 > 0, \delta_2 > 0$.
$x \to \infty$	x 趋于无穷	由无穷的邻域组成的基	$U(\infty) := \{x \in \mathbb{R} \mid \delta < \lvert x \rvert\}$, 其中 $\delta \in \mathbb{R}$.
$x \to a, x \in E$ 或 $E \ni x \to a$ 或 $x \to a \in E$	x 沿集合 E 趋于 a	点 a 在集合 E 中的去心邻域组成的基 *)	$\mathring{U}_E(a) := E \cap \mathring{U}(a)$
$x \to \infty, x \in E$ 或 $E \ni x \to \infty$ 或 $x \to \infty \in E$	x 沿集合 E 趋于无穷	在集合 E 中的无穷的邻域组成的基 **)	$U_E(\infty) := E \cap U(\infty)$

*) 假定 a 是 E 的极限点.

**) 假定集合 E 无界.

如果 $E = E_a^+ = \{x \in \mathbb{R} \mid x > a\}$ $(E = E_a^- = \{x \in \mathbb{R} \mid x < a\})$, 那么, 就把 $x \to a, x \in E$ 写成 $x \to a + 0$ $(x \to a - 0)$ 并说 x 从*右*趋于 a, 或从*大值方面*趋于

a(相应地说从左或从小值方面趋于 a). 当 $a=0$ 时, 就把 $x\to 0+0(x\to 0-0)$ 缩写成 $x\to +0(x\to -0)$.

我们还用 $E\ni x\to a+0(E\ni x\to a-0)$ 代替 $x\to a, x\in E\cap E_a^+(x\to a, x\in E\cap E_a^-)$, 它表示 x 沿集合 E 趋于 a, 但总是大于 (小于)a.

如果

$$E=E_\infty^+=\{x\in\mathbb{R}|c<x\}\quad (E=E_\infty^-=\{x\in\mathbb{R}|x<c\}),$$

就把 $x\to\infty, x\in E$ 写成 $x\to +\infty(x\to -\infty)$ 并说 *x趋于正无穷* (*x 趋于负无穷*).

用 $E\ni x\to +\infty(E\ni x\to -\infty)$ 这种写法代替 $x\to\infty, x\in E\cap E_\infty^+(x\to\infty, x\in E\cap E_\infty^-)$.

当 $E=\mathbb{N}$ 时,(如果不会发生误解) 我们就把 $x\to\infty, x\in\mathbb{N}$ 写成 $n\to\infty$, 在数列极限的理论中, 我们已经采用了这种记法.

注意, 上面举出的所有基, 都有一个共同特点, 即基中的任何两个元素之交, 还是这个基的一个元素, 而不只是包含着基的某个元素. 当我们讨论的函数不是在数轴上给出的时候, 将会遇到其他的基 [①].

还要注意, 这里所用的术语 "基", 是数学中叫做 "滤子基" 的简述, 下面所引入的对基的极限, 对于分析来说, 是现代法国数学家嘉当 (Cartan) 所创立的滤子极限的最重要的部分[②].

b. 关于基的极限

定义 12 设 $f: X\to\mathbb{R}$ 是集合 X 上的函数; $\mathcal{B}$ 是 X 中的一个基. 如果对于点 $A\in\mathbb{R}$ 的任何邻域 $V(A)$, 存在着基 $\mathcal{B}$ 中的元素 $B\in\mathcal{B}$, 使 B 的像 $f(B)$ 包含在 $V(A)$ 中, 就说 *A 是函数 $f: X\to\mathbb{R}$ 关于基 $\mathcal{B}$ 的极限*.

如果 A 是函数 $f: X\to\mathbb{R}$ 关于基 $\mathcal{B}$ 的极限, 则记

$$\lim_{\mathcal{B}} f(x)=A.$$

用逻辑符号把关于基的极限定义写出来, 就是

$$\boxed{\left(\lim_{\mathcal{B}} f(x)=A\right):=\forall V(A)\exists B\in\mathcal{B}(f(B)\subset V(A)).}$$

因为我们现在讨论数值函数, 所以把这个基本定义写成下列形式是有用处的:

$$\left(\lim_{\mathcal{B}} f(x)=A\right):=\forall\varepsilon>0\exists B\in\mathcal{B}\forall x\in B(|f(x)-A|<\varepsilon).$$

---

①例如, 平面上含有给定点的开圆 (不包含圆周) 的族是一个基. 基中的二元素之交不一定是圆, 但它总是包含着这个族里的某个圆.

②关于详细情况, 请参看Bourbaki. 一般拓扑学.

在这个说法里, 把任意邻域 $V(A)$, 换成了关于 A 点对称的邻域 (ε 邻域). 对于实函数这两种定义的等价性, 前面已经说过, 是由在一点的任何邻域中都含有该点的一个对称邻域这样一个事实推出 (请作出完整的证明).

我们已经给出了函数关于基的极限的普遍的定义. 上面又讨论了分析中最常用到的基的例子. 在具体问题中, 可能出现这些基中的某一种, 必须善于解读一般定义在具体问题中的具体含义, 并具体地写出它的基.

例如,

$$\left(\lim_{x\to a-0} f(x) = A\right) := \forall\varepsilon > 0\ \exists\delta > 0\ \forall x \in]a-\delta, a[\ (|f(x) - A| < \varepsilon),$$

$$\left(\lim_{x\to -\infty} f(x) = A\right) := \forall\varepsilon > 0\ \exists\delta \in \mathbb{R}\ \forall x < \delta\ (|f(x) - A| < \varepsilon).$$

在讨论基的例子时, 我们特别引入了无穷的邻域这一概念. 如果利用这一概念, 那么, 根据极限的一般定义, 作如下约定是合理的:

$$\left(\lim_{\mathcal{B}} f(x) = \infty\right) := \forall V(\infty)\ \exists B \in \mathcal{B}\ (f(B) \subset V(\infty))$$

同样地, 也可写成

$$\left(\lim_{\mathcal{B}} f(x) = \infty\right) := \forall\varepsilon > 0\ \exists B \in \mathcal{B}\ \forall x \in B\ (\varepsilon < |f(x)|),$$

$$\left(\lim_{\mathcal{B}} f(x) = +\infty\right) := \forall\varepsilon \in \mathbb{R}\ \exists B \in \mathcal{B}\ \forall x \in B\ (\varepsilon < f(x)),$$

$$\left(\lim_{\mathcal{B}} f(x) = -\infty\right) := \forall\varepsilon \in \mathbb{R}\ \exists B \in \mathcal{B}\ \forall x \in B\ (f(x) < \varepsilon).$$

通常, ε 总是指小的量, 但上面的定义中显然并非如此. 例如, 按我们的约定, 可以写

$$\left(\lim_{x\to +\infty} f(x) = -\infty\right) := \forall\varepsilon \in \mathbb{R}\ \exists\delta \in \mathbb{R}\ \forall x > \delta\ (f(x) < \varepsilon).$$

建议读者自己把关于有限 (数) 极限及无限极限情况下的各种基的极限定义写出来.

在第 2 段中, 我们对于特殊基 $E \ni x \to a$ 已经证明了一些有关极限的定理; 为了能够承认这些定理在普遍情况下也已得到证明, 就需在任何基下, 给出最终为常数, 最终有界与无穷小等的相应的定义.

定义 13 称函数 $f: X \to \mathbb{R}$ *对基* $\mathcal{B}$ *最终为常数*, 如果存在一数 $A \in \mathbb{R}$, 及基 $\mathcal{B}$ 中的元素 B, 使得在 B 的任何点 x 处, $f(x) = A$.

定义 14 称函数 $f: X \to \mathbb{R}$ *对基* $\mathcal{B}$ *有界*, *或对基* $\mathcal{B}$ *最终有界*, 如果存在一数 $c > 0$, 及基中的元素 $B \in \mathcal{B}$, 在每一点 $x \in B$ 处, $|f(x)| < c$.

定义 15 称函数 $f: X \to \mathbb{R}$ *对基* $\mathcal{B}$ *为无穷小*, 如果 $\lim\limits_{\mathcal{B}} f(x) = 0$.

给出了这些定义, 并观察到 "有关极限定理的证明只需要基的性质 B_1) 与 B_2)" 这个事实, 就可认为在第 2 段中所建立的所有极限性质关于任意的基全都是正确的.

特别地, 我们现在已能谈论当 $x \to \infty$, 或 $x \to -\infty$ 或 $x \to +\infty$ 时函数的极限了.

除此之外, 对于不是在数集上定义的函数, 也提供了应用极限理论的可能性, 以后将会看到, 这一点特别重要. 例如, 曲线长就是在某一曲线类上定义的一个数值函数. 如果对于折线知道了这个函数, 那么, 我们再利用极限过渡就能定义更复杂的曲线的长, 例如圆周的长.

目前, 我们从所做的这些考察和引入的与之相关的基概念所得到的主要好处在于, 不必再对每种具体形式的极限过程, 用现在的术语说, 就是对每个具体的基, 去一一验证和形式地证明有关极限的定理了.

为了完全掌握关于任意基的极限概念, 我们将在普遍形式下证明进一步的函数极限性质.

4. 函数极限的存在问题

a. 柯西准则 在叙述柯西准则之前, 先给出一个很有用的定义.

定义 16 称

$$\omega(f;E) := \sup_{x_1,x_2\in E} |f(x_1) - f(x_2)|$$

为函数 $f: X \to \mathbb{R}$ 在集合 $E \subset X$ 上的振幅. 这是在一切可能的点对 $x_1, x_2 \in E$ 上, 函数值之差的模的上确界.

例

11. $\omega(x^2;[-1,2]) = 4$.

12. $\omega(x;[-1,2]) = 3$.

13. $\omega(x;]-1,2[) = 3$.

14. $\omega(\operatorname{sgn}x;[-1,2]) = 2$.

15. $\omega(\operatorname{sgn}x;[0,2]) = 1$.

16. $\omega(\operatorname{sgn}x;]0,2[) = 0$.

定理 4 (函数极限存在的柯西准则) 设 X 为一集, $\mathcal{B}$ 为 X 中的基.

函数 $f: X \to \mathbb{R}$ 关于基有极限, 当且仅当对任何数 $\varepsilon > 0$, 存在着 $B \in \mathcal{B}$, 使得函数在 B 上的振幅小于 ε.

这样,

$$\exists \lim_{\mathcal{B}} f(x) \Leftrightarrow \forall \varepsilon > 0 \exists B \in \mathcal{B}(\omega(f;B) < \varepsilon).$$

◀ 必要性. 如果 $\lim_{\mathcal{B}} f(x) = A \in \mathbb{R}$, 那么, 对 $\varepsilon > 0$, 选出基 $\mathcal{B}$ 之元素 B, 使得在任何 $x \in B$ 处, $|f(x) - A| < \varepsilon/3$. 但这时对任何 $x_1, x_2 \in B$, 有

$$|f(x_1) - f(x_2)| < |f(x_1) - A| + |f(x_2) - A| < \frac{2\varepsilon}{3}.$$

所以, $\omega(f; B) < \varepsilon$.

充分性. 现在证准则的主要部分, 亦即, 如果对于任何 $\varepsilon > 0$, 存在基 $\mathcal{B}$ 的元素 B, 使得 $\omega(f; B) < \varepsilon$, 那么, 函数有极限.

顺次取 $1, \frac{1}{2}, \cdots, \frac{1}{n}, \cdots$ 做为 ε, 对它们取基 $\mathcal{B}$ 中的一列元素 $\tilde{B}_1, \tilde{B}_2, \cdots, \tilde{B}_n, \cdots$ 使得 $\omega(f; \tilde{B}_n) < \frac{1}{n}, n \in \mathbb{N}$. 由于 $B_n \neq \varnothing$, 可在每个 B_n 中取一个点 x_n, 则序列 $f(x_1), f(x_2), \cdots, f(x_n) \cdots$ 是基本列. 事实上, $B_n \cap B_m \neq \varnothing$, 从而借助于辅助点 $x \in B_n \cap B_m$, 就得到 $|f(x_n) - f(x_m)| \leqslant |f(x_n) - f(x)| + |f(x) - f(x_m)| < \frac{1}{n} + \frac{1}{m}$. 根据对数列已经证明了的柯西准则, $\{f_n(x), n \in \mathbb{N}\}$ 有一极限 A. 在上边的不等式中令 $m \to \infty$, 推出 $|f(x_n) - A| \leqslant \frac{1}{n}$. 由此, 并注意到 $\omega(f; B_n) < \frac{1}{n}$, 就得到了欲证的结论: 如果 $n > N = \left[\frac{2}{\varepsilon}\right] + 1$, 则在任意点 $x \in B_n$ 有 $|f(x) - A| < \varepsilon$. ▶

注 后边将看到, 上边这个证明对于在任何所谓完备空间Y 中取值的函数也是有效的, 我们现在首先感兴趣的是 $Y = \mathbb{R}$ 的情形, 这时, 可如愿地使用在序列情形证明柯西准则充分性时的那些想法.

◀ 设 $m_B = \inf_{x \in B} f(x), M_B = \sup_{x \in B} f(x)$, 并注意, 对基 $\mathcal{B}$ 的任何元 B_1, B_2 都成立 $m_{B_1} \leqslant m_{B_1 \cap B_2} \leqslant M_{B_1 \cap B_2} \leqslant M_{B_2}$, 根据完备性公理, 可以找到数 $A \in \mathbb{R}$ 分离数集 $\{m_B\}$ 和 $\{M_B\}$, 其中 $B \in \mathcal{B}$. 由于 $\omega(f; B) = M_B - m_B$, 可以断言, 只要 $\omega(f; B) < \varepsilon$, 就在每个点 $x \in B$ 成立 $|f(x) - A| < \varepsilon$. ▶

例 17 我们来证明, 当 $X = \mathbb{N}$ 且 $\mathcal{B}$ 是基 $n \to \infty, n \in \mathbb{N}$ 时, 这个函数极限存在性的一般柯西准则, 与前面讨论过的数列极限存在性的柯西准则一致.

实际上, 基 $n \to \infty, n \in \mathbb{N}$ 的元素是

$$B = \mathbb{N} \cap U(\infty) = \{n \in \mathbb{N} | N < n\},$$

它是由大于某数 $N \in \mathbb{R}$ 的所有自然数 $n \in \mathbb{N}$ 组成的集合, 可以认为 $N \in \mathbb{N}$ 而不失普遍性. 关系 $\omega(f, B) < \varepsilon$ 在现在的情况, 就是 $\forall n_1, n_2 > N$ 有 $|f(n_1) - f(n_2)| < \varepsilon$.

因此,"对于任何 $\varepsilon > 0$, 存在基 $\mathcal{B}$ 的元素 $B \in \mathcal{B}$, 使函数 $f : \mathbb{N} \to \mathbb{R}$ 在 B 上的振幅 $\omega(f, B)$ 小于 ε" 这个条件, 等价于条件 "$\{f(n)\}$ 是基本数列".

b. 复合函数的极限

定理 5 (有关复合函数极限的定理) 设 Y 是一个集合, $\mathcal{B}_Y$ 是 Y 的一个基; $g: Y \to \mathbb{R}$ 是关于基 $\mathcal{B}_Y$ 有极限的一个映射.

设 X 是一个集合, $\mathcal{B}_X$ 是 X 的一个基, 并且 $f: X \to Y$ 是 X 到 Y 中的那样的一个映射, 对于基 $\mathcal{B}_Y$ 中的任何元素 $B_Y \in \mathcal{B}_Y$, 存在着 $\mathcal{B}_X$ 的元素 $B_X \in \mathcal{B}_X$, 使得像 $f(B_X)$ 包含在 B_Y 中.

在这些条件下, 映射 f 与 g 的复合 $g \circ f: X \to \mathbb{R}$ 有定义, 关于基 $\mathcal{B}_X$ 有极限, 且

$$\lim_{\mathcal{B}_X}(g \circ f)(x) = \lim_{\mathcal{B}_Y} g(y).$$

◀ 复合函数 $g \circ f: X \to \mathbb{R}$ 有定义, 因为 $f(X) \subset Y$. 设 $\lim\limits_{\mathcal{B}_Y} g(y) = A$, 我们来证 $\lim\limits_{\mathcal{B}_Y}(g \circ f)(x) = A$. 对于 A 点的给定邻域 $V(A)$, 取基 $\mathcal{B}_Y$ 的元素 $B_Y \in \mathcal{B}_Y$, 使得 $g(B_Y) \subset V(A)$. 据题设存在基 $\mathcal{B}_X$ 的元素 $B_X \subset \mathcal{B}_X$, 使得 $f(B_X) \subset B_Y$. 这时

$$(g \circ f)(B_X) = g(f(B_X)) \subset g(B_Y) \subset V(A),$$

这样, 我们就证明了, A 是函数 $(g \circ f): X \to \mathbb{R}$ 关于基 $\mathcal{B}_X$ 的极限. ▶

例 18 $\lim\limits_{x \to 0} \dfrac{\sin 7x}{7x} = ?$

如果令 $g(y) = \dfrac{\sin y}{y}$, $f(x) = 7x$, 那么 $(g \circ f)(x) = \dfrac{\sin 7x}{7x}$. 在我们的情况, $Y = \mathbb{R} \setminus 0, X = \mathbb{R}$. 因为

$$\lim_{y \to 0} g(y) = \lim_{y \to 0} \frac{\sin y}{y} = 1,$$

因此, 为了应用定理, 必须验证, 不论取基 $y \to 0$ 的哪个元, 必能找到基 $x \to 0$ 的元素, 使它在映射 $f(x) = 7x$ 下的像, 含于事先取定的、基 $y \to 0$ 的那个元中.

基 $y \to 0$ 的元素是点 $0 \in \mathbb{R}$ 的去心邻域 $\overset{\circ}{U}_Y(0)$. 基 $x \to 0$ 的元素, 也是 $0 \in \mathbb{R}$ 点的去心邻域 $\overset{\circ}{U}_X(0)$. 令

$$\overset{\circ}{U}_Y(0) = \{y \in \mathbb{R} | \alpha < x < \beta, x \neq 0\}$$

(这里 $\alpha, \beta \in \mathbb{R}$, 且$\alpha < 0, \beta > 0$) 是 Y 的一个去心零邻域. 如果取

$$\overset{\circ}{U}_X(0) = \{x \in \mathbb{R} | \frac{\alpha}{7} < x < \frac{\beta}{7}, x \neq 0\},$$

那么, X 中的这个去心邻域就已具备了性质

$$f(\overset{\circ}{U}_X(0)) = \overset{\circ}{U}_Y(0) \subset \overset{\circ}{U}_X(0).$$

定理的条件已被满足, 所以, 可以断定

$$\lim_{x \to 0} \frac{\sin 7x}{7x} = \lim_{y \to 0} \frac{\sin y}{y} = 1.$$

例 19 函数 $g(y)=|\operatorname{sgn} y|$, 如在例 3 所见, 有极限 $\lim\limits_{y\to 0}|\operatorname{sgn} y|=1$.

函数 $y=f(x)=x\sin\frac{1}{x}$, 当 $x\neq 0$ 时有定义, 它也有极限 $\lim\limits_{x\to 0}x\sin\frac{1}{x}=0$(参看例 1)

然而函数 $(g\circ f)(x)=\left|\operatorname{sgn}\left(x\sin\frac{1}{x}\right)\right|$ 当 $x\to 0$ 时没有极限.

实际上, 在点 $x=0$ 的任何去心邻域中都有函数 $\sin\frac{1}{x}$ 的零点. 因此, 函数 $\left|\operatorname{sgn}\left(x\sin\frac{1}{x}\right)\right|$ 既能取得值 0, 又能取得值 1, 根据柯西准则, 当 $x\to 0$ 时不可能有极限.

这是否与定理相矛盾呢?

如我们在上例中所做的那样, 请检查一下, 定理的条件是否满足.

例 20 证明

$$\boxed{\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x=e.}$$

◀ 设

$Y=\mathbb{N}, \mathcal{B}_Y$ 是基 $n\to\infty, n\in\mathbb{N}$;

$X=\mathbb{R}_+=\{x\in\mathbb{R}|x>0\}, \mathcal{B}_X$ 是基 $x\to+\infty$;

$f:X\to Y$ 是映射 $x\overset{f}{\longmapsto}[x]$,

这里 $[x]$ 是数 x 的整数部分 (即不超过 x 的最大整数).

这时, 对于基 $n\to\infty, n\in\mathbb{N}$ 中任何元素 $B_Y=\{n\in\mathbb{N}|n>N\}$, 显然, 存在着基 $x\to+\infty$的元素$B_X=\{x\in\mathbb{R}|x>N+1\}$, 使得它在映射 $x\to[x]$ 下的像含于 B_Y 中.

函数 $g(n)=\left(1+\frac{1}{n}\right)^n, g_1(n)=\left(1+\frac{1}{n+1}\right)^n, g_2(n)=\left(1+\frac{1}{n}\right)^{n+1}$, 我们知道, 它们关于基 $n\to\infty, n\in\mathbb{N}$ 都有极限 e.

根据复合函数极限的定理, 可以断定函数

$$(g\circ f)(x)=\left(1+\frac{1}{[x]}\right)^{[x]},$$

$$(g_1\circ f)(x)=\left(1+\frac{1}{[x]+1}\right)^{[x]},$$

$$(g_2\circ f)(x)=\left(1+\frac{1}{[x]}\right)^{[x]+1}$$

关于基 $x\to+\infty$ 也有极限 e.

现在, 只要注意当 $x>0$ 时, 有

$$\left(1+\frac{1}{[x]+1}\right)^{[x]}<\left(1+\frac{1}{x}\right)^x<\left(1+\frac{1}{[x]}\right)^{[x]+1},$$

并且当 $x \to +\infty$ 时, 两端之项都趋于 e, 所以, 根据极限的性质 (定理 3) 即得 $\lim\limits_{x\to+\infty}\left(1+\frac{1}{x}\right)^x = e$.

现在, 我们来用复合函数极限定理证明 $\lim\limits_{x\to-\infty}\left(1+\frac{1}{x}\right)^x = e$. 我们把证明记作一列等式

$$\lim_{x\to-\infty}\left(1+\frac{1}{x}\right)^x = \lim_{(-t)\to-\infty}\left(1+\frac{1}{(-t)}\right)^{(-t)} = \lim_{t\to+\infty}\left(1-\frac{1}{t}\right)^{-t}$$
$$= \lim_{t\to+\infty}\left(1+\frac{1}{t-1}\right)^t = \lim_{t\to+\infty}\left(1+\frac{1}{t-1}\right)^{t-1}\cdot\lim_{t\to+\infty}\left(1+\frac{1}{t-1}\right)$$
$$= \lim_{t\to+\infty}\left(1+\frac{1}{t-1}\right)^{t-1} = \lim_{u\to+\infty}\left(1+\frac{1}{u}\right)^u = e.$$

这一列等式, 考虑到引用了代换 $u = t-1$ 及 $t = -x$, 归根结蒂, 它是依据复合函数极限定理得到的. 实际上, 只是当我们达到了已证得其存在性的极限 $\lim\limits_{u\to+\infty}\left(1+\frac{1}{u}\right)^u$ 时, 定理才肯定了前一个极限也存在而且等于它. 于是在这两个极限前面的那个极限也存在. 这样经有限步, 我们就到达了原来的极限. 这是复合函数极限定理应用程式在极限计算中的一个相当典型的例子.

于是, 我们得到了

$$\lim_{x\to-\infty}\left(1+\frac{1}{x}\right)^x = e = \lim_{x\to+\infty}\left(1+\frac{1}{x}\right)^x.$$

由此推出 $\lim\limits_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e$.

实际上, 设给定了 $\varepsilon > 0$.

由 $\lim\limits_{x\to-\infty}\left(1+\frac{1}{x}\right)^x = e$, 知存在 $c_1 \in \mathbb{R}$, 使得当 $x < c_1$ 时,

$$\left|\left(1+\frac{1}{x}\right)^x - e\right| < \varepsilon.$$

由 $\lim\limits_{x\to+\infty}\left(1+\frac{1}{x}\right)^x = e$, 存在 $c_2 \in \mathbb{R}$, 使得当 $c_2 < x$ 时,

$$\left|\left(1+\frac{1}{x}\right)^x - e\right| < \varepsilon,$$

于是, 当 $|x| > c = \max\{|c_1|, |c_2|\}$ 时, 就得到

$$\left|\left(1+\frac{1}{x}\right)^x - e\right| < \varepsilon.$$

这就是验证了

$$\lim_{x\to\infty}\left(1+\frac{1}{x}\right)^x = e. \blacktriangleright$$

例 21

$$\lim_{t\to 0}(1+t)^{\frac{1}{t}}=e$$

◀ 做代换 $x=1/t$ 之后, 就变成了上例中的极限了. ▶

例 22 当 $q>1$ 时, $\lim\limits_{x\to+\infty}\dfrac{x}{q^x}=0$.

◀ 已知当 $q>1$ 时 $\lim\limits_{n\to\infty}\dfrac{n}{q^n}=0$(见 §1 例 11).

现在, 像例 20 那样, 可把由函数 $[x]$(x 的整数部分) 定义的映射 $f:\mathbb{R}_+\to\mathbb{N}$ 当作辅助映射.

利用不等式

$$\frac{1}{q}\cdot\frac{[x]}{q^{[x]}}<\frac{x}{q^x}<\frac{[x]+1}{q^{[x]+1}}q,$$

并注意到, 根据复合函数极限定理, 当 $x\to\infty$ 时, 两端的极限都是零, 推知 $\lim\limits_{x\to+\infty}\dfrac{x}{q^x}=0$. ▶

例 23

$$\lim_{x\to+\infty}\frac{\log_a x}{x}=0.$$

◀ 设 $a>1$. 令 $t=\log_a x$, 则得 $x=a^t$. 根据指数函数与对数函数的性质 (注意 a^n 之无界性, $n\in\mathbb{N}$), 有 $(x\to+\infty)\Leftrightarrow(t\to+\infty)$. 利用复合函数极限定理与例 22 的结果, 得到

$$\lim_{x\to+\infty}\frac{\log_a x}{x}=\lim_{t\to+\infty}\frac{t}{a^t}=0$$

如果 $0<a<1$, 则令 $-t=\log_a x, x=a^{-t}$. 这时 $(x\to+\infty)\Leftrightarrow(t\to+\infty)$, 又因为 $1/a>1$, 所以仍然得到

$$\lim_{x\to+\infty}\frac{\log_a x}{x}=\lim_{t\to+\infty}\frac{-t}{a^{-t}}=-\lim_{t\to+\infty}\frac{t}{\left(\frac{1}{a}\right)^t}=0.$$ ▶

c. 单调函数的极限 现在讨论一类特殊但非常有用的数值函数 —— 单调函数.

定义 17 定义在数集 $E\subset\mathbb{R}$ 上的函数 $f:E\to\mathbb{R}$, 叫做在 E 上*递增的*, 如果

$$\forall x_1,x_2\in E(x_1<x_2\Rightarrow f(x_1)<f(x_2));$$

在 E 上*不降的*, 如果

$$\forall x_1,x_2\in E(x_1<x_2\Rightarrow f(x_1)\leqslant f(x_2));$$

在 E 上*不增的*, 如果

$$\forall x_1,x_2\in E(x_1<x_2\Rightarrow f(x_1)\geqslant f(x_2));$$

在 E 上递降的, 如果

$$\forall x_1, x_2 \in E (x_1 < x_2 \Rightarrow f(x_1) > f(x_2)).$$

上面这几种函数都叫做在集合 E 上的单调函数.

假定数 (或符号 $-\infty, +\infty$)$i = \inf E$ 或 $s = \sup E$ 是 E 的极限点, 且 $f: E \to \mathbb{R}$ 是 E 上的单调函数. 则以下定理成立.

定理 6 (单调函数极限存在的准则) *集合 E 上的不降函数 $f : E \to \mathbb{R}$, 当 $x \to s, x \in E$ 时有极限的充要条件是它上有界; 当 $x \to i, x \in E$ 时有极限的充要条件是它下有界.*

◀ 我们对极限 $\lim\limits_{E \ni x \to s} f(x)$ 来证明这一定理.

如果这极限存在, 则函数对基 $E \ni x \to s$ 是最终有界的.

由于 f 在 E 上不降, 可推知 f 有上界. 实际上还能断定对于任何 $x \in E, f(x) \leqslant \lim\limits_{E \ni x \to s} f(x)$. 这一点将在后文中看到.

现在转而证明, 在 f 上有界的条件下, 极限 $\lim\limits_{E \ni x \to s} f(x)$ 存在.

既然 f 上有界, 那么 f 在集合 E 上所取的值有上确界. 设 $A = \sup\limits_{x \in E} f(x)$, 今证 $\lim\limits_{E \ni x \to a} f(x) = A$. 对 $\varepsilon > 0$, 根据集合的上确界的定义, 存在点 $x_0 \in E$, 使得 $A - \varepsilon < f(x_0) \leqslant A$. 因为 f 在 E 上不降, 所以当 $x_0 < x \in E$ 时, $A - \varepsilon < f(x) \leqslant A$. 但集合 $\{x \in E | x_0 < x\}$ 显然是基 $x \to s, x \in E$ 的元素 (因为 $s = \sup E$). 于是 $\lim\limits_{E \ni x \to a} f(x) = A$ 得证.

对于极限 $\lim\limits_{E \ni x \to i} f(x)$, 可完全类似地论证. 这时,

$$\lim_{E \ni x \to i} f(x) = \inf_{x \in E} f(x). \blacktriangleright$$

d. 函数的渐近行为的比较 先用例子来说明.

设 $\pi(x)$ 是不超过给定实数 x 的素数的个数. 有那种可能性, 对于任意固定的 x 求出 $\pi(x)$ 的值 (尽管要花费过多的精力), 但我们没有能力马上回答像函数 $\pi(x)$ 当 $x \to +\infty$ 时的行为怎样, 或素数分布的渐近规律如何这样的问题. 从欧几里得时代起, 人们就已经知道, 当 $x \to +\infty$ 时, $\pi(x) \to +\infty$. 但是, $\pi(x)$ 大体上像 $\dfrac{x}{\ln x}$ 那样增长的结论, 直到 19 世纪才由切比雪夫[①]证明.

当产生需要描述函数在某一点附近 (或在无穷远处) 的行为的问题时, 通常函数本身在该点并没有定义, 这时常说, 对函数在这个点的邻域中的渐近性或渐近行为感兴趣.

---

①切比雪夫 (П. Л. Чебышев)(1821 -1894) —— 伟大的俄国数学家与力学家, 俄罗斯大数学学派的奠基人.

函数的渐近行为, 通常是用另一个较简单的或研究得比较透彻的函数来刻画, 它在所讨论的点的邻域内, 以较小的相对误差再现了所研究的函数的值.

例如, 当 $x \to +\infty$ 时, $\pi(x)$ 的行为尤如 $\frac{x}{\ln x}$; 当 $x \to 0$ 时, 函数 $\frac{\sin x}{x}$ 的行为尤如常数函数 1. 显然, 当谈及函数 $x^2 + x + \sin\frac{1}{x}$ 在 $x \to \infty$ 时的行为, 我们会说, 它基本上像函数 x^2, 而当 $x \to 0$ 时, 却如 $\sin\frac{1}{x}$.

现在, 我们给出函数渐近行为的一些基本概念的确切定义. 这些概念, 在分析学习的初期我们就将系统地应用.

定义 18 如果函数的某种性质, 或函数间的某种关系, 在基 $\mathcal{B}$ 的某个元素 $B \in \mathcal{B}$ 上成立, 就说这种性质或关系对这个基 $\mathcal{B}$ *最终被满足*.

到现在为止, 我们正是在这个意义下理解对已知基最终为常数函数, 或最终有界函数这些概念的. 我们还将在这个意义下谈论, 例如, 函数 f, g, h 最终满足关系 $f(x) = g(x)h(x)$. 这些函数甚至连原来的定义域也可能不同, 但是, 如果我们所关心的只是它们对基 $\mathcal{B}$ 的渐近行为, 那么, 对我们来说, 重要的只是它们都在 $\mathcal{B}$ 的某个元素上有定义就行了.

定义 19 我们说,*对于基 $\mathcal{B}$, 函数 f 与函数 g 比较是无穷小*, 并记作 $f \underset{\mathcal{B}}{=} o(g)$ 或对于 $\mathcal{B}, f = o(g)$, 如果关系 $f(x) = \alpha(x) \cdot g(x)$ 对基 $\mathcal{B}$ 最终被满足, 其中 $\alpha(x)$ 是对基 $\mathcal{B}$ 为无穷小函数.

例 24 当 $x \to 0$ 时 $x^2 = o(x)$, 因为 $x^2 = x \cdot x$.

例 25 当 $x \to \infty$ 时, $x = o(x^2)$, 因为最终 $x \neq 0, x = \frac{1}{x} \cdot x^2$.

由这些例子, 可得出这样的结论: 即指出 $f = o(g)$ 是在什么基之下成立这件事是完全必要的.

记号 $f = o(g)$ 的读法是 "f 等于小 og".

由定义推知, 在 $g(x) \equiv 1$ 这个特殊情况下, 记法 $f \underset{\mathcal{B}}{=} o(1)$ 所表示的恰好是对于基 $\mathcal{B}, f$ 为无穷小.

定义 20 如果 $f \underset{\mathcal{B}}{=} o(g)$ 且函数 g 本身对基 $\mathcal{B}$ 为无穷小, 就说*对于基 $\mathcal{B}, f$ 是比 g 更高阶的无穷小*.

例 26 当 $x \to \infty$ 时 $x^{-2} = \frac{1}{x^2}$ 是比 $x^{-1} = \frac{1}{x}$ 更高阶的无穷小.

定义 21 关于给定的基趋于无穷的函数叫做*对于所给基的无穷大函数*, 或简述为*对于所给基是无穷大*.

定义 22 如果 f 与 g 对于基 $\mathcal{B}$ 都是无穷大, 并且 $f \underset{\mathcal{B}}{=} o(g)$, 就说对于基 $\mathcal{B}, g$ *是比 f 更高阶的无穷大*.

例 27 当 $x \to 0$ 时, $\frac{1}{x} \to \infty$, 当 $x \to 0$ 时 $\frac{1}{x^2} \to \infty$, 且当 $x \to 0$ 时, $\frac{1}{x} = o\left(\frac{1}{x^2}\right)$. 所以, 当 $x \to 0$ 时 $\frac{1}{x^2}$ 是比 $\frac{1}{x}$ 更高阶的无穷大.

同时, 当 $x \to \infty$ 时, x^2 是比 x 更高阶的无穷大.

不要以为在选用 x^n 描述函数的渐近行为后, 就能用某个数 n——x 的幂 ——去刻画一切无穷大或无穷小.

例 28 我们来证明, 当 $a > 1$ 时, 不论对哪个 $n \in \mathbb{Z}$, 都有

$$\lim_{x \to +\infty} \frac{x^n}{a^x} = 0,$$

即当 $x \to +\infty$ 时, $x^n = o(a^x)$.

◀ 当 $n \leqslant 0$ 时, 结论显然成立. 设 $n \in \mathbb{N}$. 令 $q = \sqrt[n]{a}$, 则 $q > 1$, 且 $\frac{x^n}{a^x} = \left(\frac{x}{q^x}\right)^n$, 于是

$$\lim_{x \to +\infty} \frac{x^n}{a^x} = \lim_{x \to +\infty} \left(\frac{x}{q^x}\right)^n = \underbrace{\lim_{x \to +\infty} \frac{x}{q^x} \cdot \cdots \cdot \lim_{x \to +\infty} \frac{x}{q^x}}_{n\text{个}} = 0.$$

这里利用了归纳法原理, 乘积极限的定理以及例 22 的结果. ▶

因此, 如果 $a > 1$, 对于任何 $n \in \mathbb{Z}$, 则当 $x \to +\infty$, $x^n = o(a^x)$.

例 29 现在把上例推广, 证明当 $a > 1$ 时, 对于任何 $\alpha \in \mathbb{R}$

$$\lim_{x \to +\infty} \frac{x^\alpha}{a^x} = 0,$$

即当 $x \to +\infty$ 时, $x^\alpha = o(a^x)$.

◀ 实际上, 取数 $n \in \mathbb{N}$ 使得 $n > \alpha$, 这时, 当 $x > 1$ 时得到

$$0 < \frac{x^\alpha}{a^x} < \frac{x^n}{a^x}.$$

依据极限性质及前例之结果, 即可推得 $\lim\limits_{x \to +\infty} \frac{x^\alpha}{a^x} = 0$. ▶

例 30 证明当 $a > 1$ 时, 对任何 $\alpha \in \mathbb{R}$,

$$\lim_{\mathbb{R}_+ \ni x \to 0} \frac{a^{-\frac{1}{x}}}{x^\alpha} = 0,$$

即当 $x \to 0$ 且 $x \in \mathbb{R}_+$ 时, $a^{-\frac{1}{x}} = o(x^\alpha)$.

◀ 令 $x = \frac{1}{t}$, 据复合函数极限定理, 并利用上例之结果, 就得到

$$\lim_{\mathbb{R}_+ \ni x \to 0} \frac{a^{-\frac{1}{x}}}{x^\alpha} = \lim_{t \to +\infty} \frac{t^\alpha}{a^t} = 0. \blacktriangleright$$

例 31　证明当 $\alpha>0$ 时,

$$\lim_{x\to+\infty}\frac{\log_a x}{x^\alpha}=0,$$

即对于任何正幂指数 α, 当 $x\to+\infty$ 时, $\log_a x=o(x^\alpha)$.

◀ 如果 $a>1$, 就令 $x=a^{t/\alpha}$. 这时, 根据指数函数与对数函数的性质, 复合函数极限定理, 以及例 29 之结果, 求得

$$\lim_{x\to+\infty}\frac{\log_a x}{x^\alpha}=\lim_{t\to+\infty}\frac{\left(\dfrac{t}{\alpha}\right)}{a^t}=\alpha^{-1}\cdot\lim_{t\to+\infty}\frac{t}{a^t}=0.$$

如果 $0<a<1$, 则 $\dfrac{1}{a}>1$, 作代换 $x=a^{-t/\alpha}$ 后得到

$$\lim_{x\to+\infty}\frac{\log_a x}{x^\alpha}=\lim_{t\to+\infty}\frac{\left(-\dfrac{t}{\alpha}\right)}{a^{-t}}=-\frac{1}{\alpha}\lim_{t\to+\infty}\frac{t}{\left(\dfrac{1}{a}\right)^t}=0.$$ ▶

例 32　再证明对于任何 $\alpha>0$,

$$x^\alpha\log_a x=o(1), \text{当} x\to 0, x\in\mathbb{R}_+.$$

◀ 只要证当 $\alpha>0$ 时 $\lim\limits_{\mathbb{R}_+\ni x\to 0}x^\alpha\log_a x=0$. 设 $x=\dfrac{1}{t}$, 应用复合函数极限定理及前例之结果, 得到

$$\lim_{\mathbb{R}_+\ni x\to 0}x^\alpha\log_a x=\lim_{t\to+\infty}\frac{\log_a\left(\dfrac{1}{t}\right)}{t^\alpha}=-\lim_{t\to+\infty}\frac{\log_a t}{t^\alpha}=0.$$ ▶

定义 23　我们约定, 把两个函数 f,g 对于基 $\mathcal{B}$ 最终满足关系 $f(a)=\beta(x)g(x)$ 记作 $f\underset{\mathcal{B}}{=}O(g)$ 或对于基 $\mathcal{B}, f=O(g)$,(读作对于 $\mathcal{B}, f$ 等于大Og), 其中的 $\beta(x)$ 是关于 $\mathcal{B}$ 最终有界的函数.

特别地, 记号 $f\underset{\mathcal{B}}{=}O(1)$ 表示对于基 $\mathcal{B}$, 函数 f 最终有界.

例 33　当 $x\to\infty$ 时, $\left(\dfrac{1}{x}+\sin x\right)x=O(x)$.

定义 24　如果 $f\underset{\mathcal{B}}{=}O(g)$ 且 $g\underset{\mathcal{B}}{=}O(f)$, 就说对于基 $\mathcal{B}$, 函数 f 与 g 是同阶的, 并记作: 对于基 $\mathcal{B}, f\asymp g$.

例 34　当 $x\to\infty$ 时, 函数 $(2+\sin x)x$ 与 x 是同阶的, 但当 $x\to\infty$ 时, $(1+\sin x)x$ 与 x 不同阶.

条件：f 与 g 对于基 $\mathcal{B}$ 同阶，显然等价于：存在数 $c_1>0, c_2>0$ 及基 $\mathcal{B}$ 之元素 B，使得在 B 上，关系式

$$c_1|g(x)| \leqslant |f(x)| \leqslant c_2|g(x)|$$

或

$$\frac{1}{c_2}|f(x)| \leqslant |g(x)| \leqslant \frac{1}{c_1}|f(x)|$$

成立.

定义 25 如果函数 f 与 g 对于基 $\mathcal{B}$ 最终满足关系 $f(x)=\gamma(x)g(x)$，其中 $\lim\limits_{\mathcal{B}}\gamma(x)=1$，就说对于基 $\mathcal{B}$，函数 f 渐近于函数 g，或者简称为 f 与 g 等价.

我们把 f 与 g 的这种关系记成 $f \underset{\mathcal{B}}{\sim} g$，或对于基 $\mathcal{B}$，$f \sim g$.

我们用"等价"这个词是合适的，因为

$$(f \underset{\mathcal{B}}{\sim} f),$$
$$(f \underset{\mathcal{B}}{\sim} g) \Rightarrow (g \underset{\mathcal{B}}{\sim} f),$$
$$(f \underset{\mathcal{B}}{\sim} g) \wedge (g \underset{\mathcal{B}}{\sim} h) \Rightarrow (f \underset{\mathcal{B}}{\sim} h).$$

实际上，关系 $f \underset{\mathcal{B}}{\sim} f$ 是显然的，这时 $\gamma(x)\equiv 1$. 此外，如果 $\lim\limits_{\mathcal{B}}\gamma(x)=1$，那么，

$$\lim_{\mathcal{B}}\frac{1}{\gamma(x)}=1, \text{且} g(x)=\frac{1}{\gamma(x)}f(x).$$

这里只需说明，为什么可以认为 $\gamma(x)\neq 0$. 如果在元素 $B_1\in\mathcal{B}$ 上 $f(x)=\gamma(x)g(x)$ 成立，而在元素 $B_2\in\mathcal{B}$ 上 $\frac{1}{2}<|\gamma(x)|<\frac{3}{2}$ 成立，那么，就能在 $\mathcal{B}$ 中取元素 $B\subset B_1\cap B_2$，使以上两关系都成立，在 B 之外，为了方便，可认为处处 $\gamma(x)\equiv 1$. 这样一来，确实 $(f\sim g)\Rightarrow(g\sim f)$.

最后，如果在 $B_1\in\mathcal{B}$ 上 $f(x)=\gamma_1(x)g(x)$，而在 $B_2\in\mathcal{B}$ 上 $g(x)=\gamma_2(x)h(x)$，那么取 $B\subset B_1\cap B_2, B\in\mathcal{B}$，则在 B 上以上二关系同时成立，所以在 B 上 $f(x)=\gamma_1(x)\gamma_2(x)h(x)$. 但

$$\lim_{\mathcal{B}}\gamma_1(x)\gamma_2(x)=\lim_{\mathcal{B}}\gamma_1(x)\cdot\lim_{\mathcal{B}}\gamma_2(x)=1;$$

这样就验证了 $f \underset{\mathcal{B}}{\sim} h$.

指出以下的事实是有好处的：因为 $\lim\limits_{\mathcal{B}}\gamma(x)=1$ 与 $\gamma(x)=1+\alpha(x)$ 等价，其中 $\lim\limits_{\mathcal{B}}\alpha(x)=0$，所以对于基 $\mathcal{B}$，$f \underset{\mathcal{B}}{\sim} g$ 与 $f(x)=g(x)+\alpha(x)g(x)=g(x)+o(g(x))$ 等价.

我们看到，若函数 $g(x)$ 与 $f(x)$ 对于基 $\mathcal{B}$ 等价，则借助函数 $g(x)$ 去逼近函数 $f(x)$ 时，其相对误差 $|\alpha(x)|=\left|\frac{f(x)-g(x)}{g(x)}\right|$ 是对于 $\mathcal{B}$ 的无穷小.

我们来讨论几个例题.

例 35　当 $x \to \infty$ 时, $x^2 + x = \left(1 + \frac{1}{x}\right) x^2 \sim x^2$,

此二函数之差的绝对值

$$|(x^2 + x) - x^2| = |x|$$

趋于无穷. 当 $x^2 + x$ 用与之等价的 x^2 代替时, 相对误差 $\frac{|x|}{x^2} = \frac{1}{|x|}$ 当 $x \to \infty$ 时, 趋于零.

例 36　在本段开头曾谈到著名的素数分布的渐近规律. 现在我们已能写出它的精确表达式了:

$$\pi(x) = \frac{x}{\ln x} + o\left(\frac{x}{\ln x}\right), \text{当} x \to \infty \text{时}.$$

例 37　因为 $\lim\limits_{x \to 0} \frac{\sin x}{x} = 1$, 所以当 $x \to 0$ 时, $\sin x \sim x$, 这又可以写成等式 $\sin x = x + o(x)$ 当 $x \to 0$ 的形式.

例 38　证明当 $x \to 0$ 时, $\ln(1 + x) \sim x$.

◀ $\lim\limits_{x \to 0} \frac{\ln(1+x)}{x} = \lim\limits_{x \to 0} \ln(1+x)^{\frac{1}{x}} = \ln\left(\lim\limits_{x \to 0}(1+x)^{\frac{1}{x}}\right) = \ln e = 1.$

在第一个等式中用了 $\log_a(b^\alpha) = \alpha \log_a b$, 在第二个等式中用了 $\lim\limits_{t \to b} \log_a t = \log_a b = \log_a \left(\lim\limits_{t \to b} t\right)$. ▶

于是, 当 $x \to 0$ 时, $\ln(1 + x) = x + o(x)$.

例 39　证明当 $x \to 0$ 时, $e^x = 1 + x + o(x)$.

◀ $\lim\limits_{x \to 0} \frac{e^x - 1}{x} = \lim\limits_{t \to 0} \frac{t}{\ln(1+t)} = 1.$

在上式中, 我们作了代换 $x = \ln(1 + t), e^x - 1 = t$, 并利用了当 $x \to 0$ 时, $e^x \to e^0 = 1$, 以及当 $x \neq 0$ 时, $e^x \neq 1$. 这样, 根据复合极限定理及上例之结果, 即得欲证. ▶

于是, 当 $x \to 0$ 时, $e^x - 1 \sim x$.

例 40　证明当 $x \to 0$ 时, $(1 + x)^\alpha = 1 + \alpha x + o(x)$.

$$◀ \lim_{x \to 0} \frac{(1+x)^\alpha - 1}{x} = \lim_{x \to 0} \frac{e^{\alpha \ln(1+x)} - 1}{\alpha \ln(1+x)} \cdot \frac{\alpha \ln(1+x)}{x} = \alpha \lim_{t \to 0} \frac{e^t - 1}{t} \cdot \lim_{x \to 0} \frac{\ln(1+x)}{x} = \alpha.$$

在计算时, 假定了 $\alpha \neq 0$, 做了代换 $\alpha \ln(1 + x) = t$, 并利用了前两例之结果. 如果 $\alpha = 0$ 时, 结论是显然的. ▶

于是, 当 $x \to 0$ 时, $(1 + x)^\alpha - 1 \sim \alpha x$.

当计算极限时, 有时要用到下面的简单命题.

命题 3 设 $f \underset{B}{\sim} \widetilde{f}$, 那么, $\lim\limits_{B} f(x)g(x) = \lim\limits_{B} \widetilde{f}(x)g(x)$, 如果这两个极限至少有一个存在.

◀ 实际上, 既然 $f(x) = \gamma(x)\widetilde{f}(x)$ 且 $\lim\limits_{B} \gamma(x) = 1$, 就有,

$$\begin{aligned}\lim_{B} f(x)g(x) &= \lim_{B} \gamma(x)\widetilde{f}(x)g(x) \\ &= \lim_{B} \gamma(x) \lim_{B} \widetilde{f}(x)g(x) = \lim_{B} \widetilde{f}(x)g(x). \blacktriangleright\end{aligned}$$

例 41

$$\begin{aligned}\lim_{x\to 0} \frac{\ln \cos x}{\sin x^2} &= \frac{1}{2} \lim_{x\to 0} \frac{\ln \cos^2 x}{x^2} = \frac{1}{2} \lim_{x\to 0} \frac{\ln(1-\sin^2 x)}{x^2} \\ &= \frac{1}{2} \lim_{x\to 0} \frac{-\sin^2 x}{x^2} = -\frac{1}{2} \lim_{x\to 0} \frac{x^2}{x^2} = -\frac{1}{2}.\end{aligned}$$

在这里, 我们利用了: 当 $\alpha \to 0$ 时 $\ln(1+\alpha) \sim \alpha$; 当 $x \to 0$ 时 $\sin x \sim x$; 当 $\beta \to 0$ 时 $\dfrac{1}{\sin \beta} \sim \dfrac{1}{\beta}$, 以及当 $x \to 0$ 时, $\sin^2 x \sim x^2$.

我们证明了, 在求单项式极限时, 对于已知基相互等价函数可彼此替代. 但这条规则不能应用于求函数的和或差.

例 42 当 $x \to +\infty$ 时, $\sqrt{x^2+x} \sim x$, 但是,

$$\lim_{x\to\infty} (\sqrt{x^2+x} - x) \neq \lim_{x\to\infty} (x - x) = 0.$$

实际上,

$$\lim_{x\to\infty} (\sqrt{x^2+x} - x) = \lim_{x\to+\infty} \frac{x}{\sqrt{x^2+x}+x} = \lim_{x\to\infty} \frac{1}{\sqrt{1+\frac{1}{x}}+1} = \frac{1}{2}.$$

我们还应注意关于记号 $o(\ )$, $O(\ )$ 在分析中广泛用到的几个变换规则.

命题 4 对于给定的基, 有

a) $o(f) + o(f) = o(f)$;

b) $o(f)$ 也是 $O(f)$;

c) $o(f) + O(f) = O(f)$;

d) $O(f) + O(f) = O(f)$;

e) 如果 $g(x) \neq 0$, 那么

$$\frac{o(f(x))}{g(x)} = o\left(\frac{f(x)}{g(x)}\right),$$

并且

$$\frac{O(f(x))}{g(x)} = O\left(\frac{f(x)}{g(x)}\right).$$

请注意由符号 $o(\cdot)$ 和 $O(\cdot)$ 本身的含义而来的它们的运算的特殊性. 例如, $2o(f)$ $o(f)$, 或 $o(f)+O(f)=O(f)$(虽然, 一般说来, $o(f)\neq 0$), 或 $o(f)=O(f)$, 但 $O(f)\neq o(f)$. 这里等式的等号有词 "是" 的意思. 符号 $o(\cdot)$ 和 $O(\cdot)$ 本身所表示的与其说是函数, 倒不如说是函数的渐近行为的特征, 正好这个特征是许多函数具有的. 例如, f 有, $2f$ 也有, 等等.

◀ a) 在明确了上边说的这些事情以后, 就不会觉得命题 a) 太突然了, 它的第一个符号 $o(f)$ 表示某个形如 $\alpha_1(x)f(x)$ 的函数, 这里 $\lim\limits_{B}\alpha_1(x)=0$. 第二个符号 (如果必要, 可给它加上记号以示区别于第一个符号 $o(f)$) 表示某个形如 $\alpha_2(x)f(x)$ 的函数, 其中 $\lim\limits_{B}\alpha_2(x)=0$.

因此,

$$\alpha_1(x)f(x)+\alpha_2(x)f(x)=(\alpha_1(x)+\alpha_2(x))f(x)=\alpha_3(x)f(x),$$

这里, $\lim\limits_{B}\alpha_3(x)=0$.

b) 可从有极限的函数必定最终有界推知.

c) 由 b) 与 d) 推知.

d) 由最终有界函数之和仍为最终有界推知.

e) $\dfrac{o(f(x))}{g(x)}=\dfrac{\alpha(x)f(x)}{g(x)}=\alpha(x)\cdot\dfrac{f(x)}{g(x)}=o\left(\dfrac{f(x)}{g(x)}\right)$

可类似地验证 e) 的第二部分. ▶

现在, 利用这些规则及例 40 中得到的等价关系, 就能用以下方法直接求出例 42 中的极限:

$$\begin{aligned}\lim_{x\to+\infty}(\sqrt{x^2+x}-x)&=\lim_{x\to+\infty}x\left(\sqrt{1+\frac{1}{x}}-1\right)\\&=\lim_{x\to+\infty}x\left(1+\frac{1}{2}\frac{1}{x}+o\left(\frac{1}{x}\right)-1\right)=\lim_{x\to+\infty}\left(\frac{1}{2}+x\cdot o\left(\frac{1}{x}\right)\right)\\&=\lim_{x\to+\infty}\left(\frac{1}{2}+o(1)\right)=\frac{1}{2}.\end{aligned}$$

不久, 我们将证明下面几个重要公式, 现在, 这些公式已经值得我们像九九表那样熟记:

$$e^x=1+\frac{1}{1!}x+\frac{1}{2!}x^2+\cdots+\frac{1}{n!}x^n+\cdots, \text{当}x\in\mathbb{R}.$$

$$\cos x=1-\frac{1}{2!}x^2+\frac{1}{4!}x^4+\cdots+\frac{(-1)^k}{(2k)!}x^{2k}+\cdots, \text{当}x\in\mathbb{R}.$$

$$\sin x=\frac{1}{1!}x-\frac{1}{3!}x^3+\cdots+\frac{(-1)^k}{(2k+1)!}x^{2k+1}+\cdots, \text{当}x\in\mathbb{R}$$

$$\ln(1+x)=x-\frac{1}{2}x^2+\frac{1}{3}x^3+\cdots+\frac{(-1)^{n-1}}{n}x^n+\cdots, \text{当}|x|<1,$$

$$(1+x)^\alpha=1+\frac{\alpha}{1!}x+\frac{\alpha(\alpha-1)}{2!}x^2+\cdots+\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n+\cdots, \text{当}|x|<1.$$

这些关系式, 一方面可当作计算公式, 另一方面, 即将看到, 它们本身就包含了以下的渐近公式, 这些渐近公式是例 37 40 中所得公式的推广:

$$e^x = 1 + \frac{1}{1!}x + \cdots + \frac{1}{n!}x^n + O(x^{n+1}), \text{当} x \to 0$$

$$\cos x = 1 - \frac{1}{2!}x^2 + \cdots + \frac{(-1)^k}{(2k)!}x^{2k} + O(x^{2k+2}), \text{当} x \to 0$$

$$\sin x = \frac{1}{1!}x - \frac{1}{3!}x^3 + \cdots + \frac{(-1)^k}{(2k+1)!}x^{2k+1} + O(x^{2k+3}), \text{当} x \to 0$$

$$\ln(1+x) = x - \frac{1}{2}x^2 + \cdots + \frac{(-1)^{n-1}}{n}x^n + O(x^{n+1}), \text{当} x \to 0$$

$$(1+x)^\alpha = 1 + \frac{\alpha}{1!}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n + O(x^{n+1}), \quad \text{当} x \to 0.$$

这些公式通常是求初等函数极限的最有效的工具. 同时指出一个有用的事实: 当 $x \to 0$ 时,

$$O(x^{m+1}) = x^{m+1} \cdot O(1) = x \cdot x^m \cdot O(1) = x^m \cdot o(1) = o(x^m).$$

最后, 作为应用这些公式的示范, 讨论几个例子.

例 43

$$\lim_{x\to0} \frac{x - \sin x}{x^3} = \lim_{x\to0} \frac{x - \left(x - \frac{1}{3!}x^3 + O(x^5)\right)}{x^3} = \lim_{x\to0} \left(\frac{1}{3!} + O(x^2)\right) = \frac{1}{3!}.$$

例 44 $\lim\limits_{x\to\infty} x^2 \left(\sqrt[7]{\dfrac{x^3+x}{1+x^3}} - \cos\dfrac{1}{x}\right) = ?$

当 $x \to \infty$, 有:

$$\frac{x^3+x}{1+x^3} = \frac{1+\frac{1}{x^2}}{1+\frac{1}{x^3}} = \left(1+\frac{1}{x^2}\right)\left(1+\frac{1}{x^3}\right)^{-1}$$
$$= \left(1+\frac{1}{x^2}\right)\left(1-\frac{1}{x^3}+O\left(\frac{1}{x^6}\right)\right) = 1 + \frac{1}{x^2} + O\left(\frac{1}{x^3}\right),$$

$$\sqrt[7]{\frac{x^3+x}{1+x^3}} = \left(1+\frac{1}{x^2}+O\left(\frac{1}{x^3}\right)\right)^{\frac{1}{7}} = 1 + \frac{1}{7}\frac{1}{x^2} + O\left(\frac{1}{x^3}\right),$$

$$\cos\frac{1}{x} = 1 - \frac{1}{2!}\frac{1}{x^2} + O\left(\frac{1}{x^4}\right),$$

由此得到

$$\sqrt[7]{\frac{x^3+x}{1+x^3}}-\cos\frac{1}{x}=\frac{9}{14}\frac{1}{x^2}+O\left(\frac{1}{x^3}\right),\text{当}x\to\infty.$$

因此, 所求之极限等于

$$\lim_{x\to\infty}x^2\left(\frac{9}{14x^2}+O\left(\frac{1}{x^3}\right)\right)=\frac{9}{14}.$$

例 45

$$\begin{aligned}\lim_{x\to\infty}\left[\frac{\left(1+\frac{1}{x}\right)^x}{e}\right]^x &= \lim_{x\to\infty}\exp\left\{x\left[\ln\left(1+\frac{1}{x}\right)^x-1\right]\right\}\\ &= \lim_{x\to\infty}\exp\left\{x^2\ln\left(1+\frac{1}{x}\right)-x\right\}\\ &= \lim_{x\to\infty}\exp\left\{x^2\left[\frac{1}{x}-\frac{1}{2x^2}+O\left(\frac{1}{x^3}\right)\right]-x\right\}\\ &= \lim_{x\to\infty}\exp\left\{-\frac{1}{2}+O\left(\frac{1}{x}\right)\right\}=e^{-\frac{1}{2}}.\end{aligned}$$

练 习

1. a) 试证, 定义在 $\mathbb{R}$ 上且满足要求

$$\begin{gathered}f(1)=a,a>0,a\neq 0,\\ f(x_1)\cdot f(x_2)=f(x_1+x_2),\\ \text{当}x\to x_0\text{时},f(x)\longrightarrow f(x_0)\end{gathered}$$

的函数存在且唯一.

b) 试证, 在 $\mathbb{R}_+$ 上定义且满足条件

$$\begin{aligned}f(a) &= 1,\\ f(x_1)+f(x_2) &= f(x_1\cdot x_2),\end{aligned}$$

当 $x_0\in\mathbb{R}_+$ 且 $\mathbb{R}_+\ni x\to x_0$ 时, $f(x)\to f(x_0)$的函数存在且唯一.

提示 再次研究例 10 中所讨论的指数函数与对数函数的结构.

2. a) 试构造双方单值对应 $\varphi:\mathbb{R}\to\mathbb{R}_+$ 满足: 对任何 $x,y\in\mathbb{R}$ 成立 $\varphi(x+y)=\varphi(x)\cdot\varphi(y)$, 亦即, 原像 (在 $\mathbb{R}$ 中) 的加法运算对应像 (在 $\mathbb{R}_+$ 中) 的乘法运算. 这样的映射的存在性表明, 群 $(\mathbb{R},+)$ 和 $(\mathbb{R}_+,\cdot)$ 作为代数对象是一样的, 或者说, 它们同构.

b) 试证, 群 $(\mathbb{R},+)$ 与 $(\mathbb{R}\setminus 0_+,\cdot)$ 不同构.

3. 求极限:

a) $\lim\limits_{x\to+0} x^x$;

b) $\lim\limits_{x\to+\infty} x^{\frac{1}{x}}$;

c) $\lim\limits_{x\to0} \dfrac{\log_a(1+x)}{x}$;

d) $\lim\limits_{x\to0} \dfrac{a^x-1}{x}$.

4. 试证

$$1+\frac{1}{2}+\cdots+\frac{1}{n}=\ln n+c+o(1), \text{当} n\to\infty.$$

这里 c 是常数 ($c=0.577\ 21\cdots$ 叫做欧拉常数).

提示 可用

$$\ln\frac{n+1}{n}=\ln\left(1+\frac{1}{n}\right)=\frac{1}{n}+O\left(\frac{1}{n^2}\right), \text{当} n\to\infty.$$

5. 试证

a) 如果两个级数 $\sum\limits_{n=1}^{\infty} a_n$, $\sum\limits_{n=1}^{\infty} b_n$ 都是正项级数, 且当 $n\to\infty$ 时 $a_n\sim b_n$, 则这两个级数同时收敛或同时发散.

b) 级数 $\sum\limits_{n=1}^{\infty} \sin\dfrac{1}{n^p}$ 仅当 $p>1$ 时收敛.

6. 试证

a) 如果对于任何 $n\in\mathbb{N}$ 有 $a_n\geqslant a_{n+1}>0$, 且 $\sum\limits_{n=1}^{\infty} a_n$ 收敛, 则当 $n\to\infty$ 时 $a_n=o\left(\dfrac{1}{n}\right)$;

b) 如果 $b_n=o\left(\dfrac{1}{n}\right)$, 则必能做出一个收敛级数 $\sum\limits_{n=1}^{\infty} a_n$, 使得当 $n\to\infty$ 时, $b_n=o(a_n)$;

c) 如果正项级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛, 则以 $A_n=\sqrt{\sum\limits_{k=n-1}^{\infty} a_k}-\sqrt{\sum\limits_{k=n}^{\infty} a_k}$ 为项的级数 $\sum\limits_{n=2}^{\infty} A_n$ 也收敛, 并且当 $n\to\infty$ 时, $a_n=o(A_n)$;

d) 如果正项级数 $\sum\limits_{n=1}^{\infty} a_n$ 不收敛, 那么, 以 $A_n=\sqrt{\sum\limits_{k=1}^{n} a_k}-\sqrt{\sum\limits_{k=1}^{n-1} a_k}$ 为项的级数 $\sum\limits_{n=2}^{\infty} A_n$ 也不收敛, 并且当 $n\to\infty$ 时, $A_n=o(a_n)$.

由 c) 与 d) 推知, 没有任何收敛 (发散) 级数, 能够做为用比较法对其他级数判断收敛 (发散) 的万能的标准级数.

7. 试证

a) 级数 $\sum\limits_{n=1}^{\infty} \ln a_n, a_n>0, n\in\mathbb{N}$, 收敛, 当且仅当数列 $\{\Pi_n=a_1\cdot\cdots\cdot a_n\}$ 有不为零的极限;

b) 级数 $\sum\limits_{n=1}^{\infty} \ln(1+a_n)$ 绝对收敛, 当且仅当级数 $\sum\limits_{n=1}^{\infty} a_n$ 绝对收敛, 其中 $|a_n|<1$.

提示 参看问题 5a).

8. 设数列 $\Pi_n=\prod\limits_{k=1}^{n} e_k$ 有有限的异于零的极限 Π, 就说无穷乘积 $\prod\limits_{k=1}^{\infty} e_k$ 收敛, 且记 $\prod\limits_{k=1}^{\infty} e_k=\Pi$.

试证

a) 如果无穷乘积 $\prod\limits_{n=1}^{\infty} e_n$ 收敛, 则当 $n\to\infty$ 时, $e_n\to1$;

b) 如果 $\forall n\in\mathbb{N}(e_n>0)$, 那么, 无穷乘积 $\prod\limits_{n=1}^{\infty} e_n$ 收敛, 当且仅当 $\sum\limits_{n=1}^{\infty} \ln e_n$ 收敛;

c) 如果 $e_n = 1 + a_n$, 并且所有 a_n 同号, 那么, 无穷乘积 $\prod\limits_{n=1}^{\infty}(1+a_n)$ 收敛, 当且仅当级数 $\sum\limits_{n=1}^{\infty} a_n$ 收敛.

9. 试求

a) $\prod\limits_{n=1}^{\infty}(1+x^{2n-1})$;

b) $\prod\limits_{n=1}^{\infty}\cos\frac{x}{2^n}$, 并证明韦达[①]公式

$$\frac{\pi}{2} = \frac{1}{\sqrt{\frac{1}{2}}\cdot\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}}}\cdot\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}+\frac{1}{2}\sqrt{\frac{1}{2}}}}\cdots}$$

c) 函数 $f(x)$, 如果

$$f(0) = 1,$$
$$f(2x) = \cos^2 x \cdot f(x),$$
$$当x \to 0时, f(x) \to f(0).$$

提示　$x = 2\cdot\frac{x}{2}$.

10. 试证

a) 如果 $\frac{b_n}{b_{n+1}} = 1 + \beta_n (n = 1, 2, \cdots)$, 并且级数 $\sum\limits_{n=1}^{\infty}\beta_n$ 绝对收敛, 则 $\exists \lim\limits_{n\to\infty} b_n = b \in \mathbb{R}$;

b) 如果 $\frac{a_n}{a_{n+1}} = 1 + \frac{p}{n} + \alpha_n (n = 1, 2, \cdots)$, 并且级数 $\sum\limits_{n=1}^{\infty}\alpha_n$ 绝对收敛, 则 $a_n \sim \frac{c}{n^p}$, 当 $n \to \infty$;

c) 如果级数 $\sum\limits_{n=1}^{\infty} a_n$ 满足 $\frac{a_n}{a_{n+1}} = 1 + \frac{p}{n} + \alpha_n$, 且级数 $\sum\limits_{n=1}^{\infty}\alpha_n$ 绝对收敛, 则级数 $\sum\limits_{n=1}^{\infty} a_n$ 当 $p > 1$ 时绝对收敛, 而当 $p \leqslant 1$ 时发散 (*级数绝对收敛性的高斯检验法*).

11. 试证, 对任何正项序列 $\{a_n\}$ 必有

$$\varlimsup_{n\to\infty}\left(\frac{1+a_{n+1}}{a_n}\right)^n \geqslant e,$$

且这一估值是最好的.

---

① 韦达 (Veite)(1540—1603)　— 法国数学家, 现代代数符号体系的创始人之一.

第四章　连续函数

§1. 基本定义和例子

1. 函数在一点处的连续性　设 f 是定义在点 $a \in \mathbb{R}$ 的某个邻域中的实值函数.

直观地说, 函数 f 在点 a 处连续, 是指它的值 $f(x)$ 随着自变量 x 趋近点 a 而趋近于函数在该点 a 处的值 $f(a)$.

现在我们把函数在一点连续概念的这种描述精确化.

定义 0　称函数 f 在点 a 处连续, 如果对于函数在点 a 处的值 $f(a)$ 的任何一个邻域 $V(f(a))$, 都存在点 a 的这样一个邻域 $U(a)$, 它在映射 f 下的像含在 $V(f(a))$ 中.

我们引入这个定义的形式逻辑写法及其在分析学中常用的两个变形:

(f 在点 a 连续)$:= (\forall V(f(a))\ \exists U(a)\ (f(U(a)) \subset V(f(a))))$,

$\forall \varepsilon > 0\ \exists U(a)\ \forall x \in U(a)\ (|f(x) - f(a)| < \varepsilon)$,

$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in \mathbb{R}\ (|x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon)$.

对于实值函数, 这几种表述是等价的, 这是由于 (正像已经不止一次地指出过的那样) 点的任何一个邻域都包含该点的某个对称邻域.

譬如说, 如果对于点 $f(a)$ 的任意一个 ε 邻域 $V^\varepsilon(f(a))$ 都可以选出点 a 的一个邻域 $U(a)$, 使得

$$\forall x \in U(a)(|f(x) - f(a)| < \varepsilon) \text{即} f(U(a)) \subset V^\varepsilon(f(a))$$

的话, 那么, 对于任意一个邻域 $V(f(a))$, 同样可以选出点 a 的相应邻域来. 实际上, 只要先取 $V^\varepsilon(f(a))\subset V(f(a))$, 然后再根据 $V^\varepsilon(f(a))$ 来找 $U(a)$. 于是

$$f(U(a))\subset V^\varepsilon(f(a))\subset V(f(a)).$$

这样一来, 如果函数在点 a 依上述第二种定义是连续的, 则它依原始的定义亦在此点连续. 反过来是明显的. 因此前两种表述的等价性已经得到了验证.

进一步的验证留给读者来做.

为了把注意力集中于函数在一点处的连续性这一基本概念, 起初我们简单地假定了函数 f 在点 a 的整个邻域上有定义. 现在来考察一般的情形.

设 $f:E\to\mathbb{R}$ 是定义在某个集合 $E\subset\mathbb{R}$ 上的实值函数, 而 a 是函数定义域中的一点.

定义 1 称函数 $f:E\to\mathbb{R}$ 在点 $a\in E$ *连续*, 如果对于函数在 a 点所取的值 $f(a)$ 的任意邻域 $V(f(a))$, 存在点 a 在集 E 中的一个邻域 $U_E(a)$[①], 它的像 $f(U_E(a))$ 含在 $V(f(a))$ 中.

于是

$$\boxed{(f:E\to\mathbb{R}\text{在}a\in E\text{连续}):=(\forall V(f(a))\ \exists U_E(a)\ (f(U_E(a))\subset V(f(a)))).}$$

当然, 定义 1 也可以用前面见到的那样的 $\varepsilon-\delta$ 形式写出. 在需要做数值估计的情况, 这种形式的定义常常是有用的, 甚至是必需的.

我们写出定义 1 的这些变形:

$$(f:E\to\mathbb{R}\text{在}a\in E\text{连续}):=(\forall\varepsilon>0\ \exists U_E(a)\ \forall x\in U_E(a)\ (|f(x)-f(a)|<\varepsilon)).$$

或者

$$(f:E\to\mathbb{R}\text{在}a\in E\text{连续}):=(\forall\varepsilon>0\ \exists\delta>0\ \forall x\in E\ (|x-a|<\delta\Rightarrow|f(x)-f(a)|<\varepsilon)).$$

现在我们来详细地讨论函数在一点处连续的概念.

1° 如果 a 是集 E 的孤立点, 也就是说不是 E 的极限点, 那么, 存在 a 点的一个邻域 $U(a)$, 其中除了 a 点本身外不含有 E 的别的点. 在这种情况下, $U_E(a)=a$, 因此不管邻域 $V(f(a))$ 如何, 恒有 $f(U_E(a))=f(a)\subset V(f(a))$. 于是, 函数在定义域的任何孤立点处显然都是连续的. 但这是一种蜕化的情形.

2° 因此, 连续这一概念的有实际内容的情形是 $a\in E$ 且 a 是 E 的极限点的情形. 从定义 1 看到

$$(f:E\to\mathbb{R}\text{在}E\text{的极限点}a\in E\text{处连续})\Leftrightarrow(\lim_{E\ni x\to a}f(x)=f(a)).$$

---

①注意, $U_E(a)=E\cap U(a)$

◀ 事实上, 如果 a 是 E 的极限点, 则点 a 的空心邻域 $\mathring{U}_E(a) = U_E(a) \backslash a$ 的基 $E \ni x \to a$ 有定义.

若 f 在 a 连续, 那么, 对于邻域 $V(f(a))$, 找一个邻域 $U_E(a)$ 使 $f(U_E(a)) \subset V(f(a))$, 同时就有 $f(\mathring{U}_E(a)) \subset V(f(a))$, 于是根据极限的定义, $\lim\limits_{E \ni x \to a} f(x) = f(a)$.

反之, 如果知道 $\lim\limits_{E \ni x \to a} f(x) = f(a)$, 则对于邻域 $V(f(a))$, 存在空心邻域 $\mathring{U}_E(a)$ 使 $f(\mathring{U}_E(a)) \subset V(f(a))$. 但由于 $f(a) \in V(f(a))$, 所以 $f(U_E(a)) \subset V(f(a))$. 根据定义 1, 这就意味着函数 f 在点 $a \in E$ 连续. ▶

3° 由于关系式 $\lim\limits_{E \ni x \to a} f(x) = f(a)$ 可以改写成

$$\boxed{\lim_{E \ni x \to a} f(x) = f(\lim_{E \ni x \to a} x),}$$

我们现在得到这样一个有用的结论: 在一点连续的函数或运算, 并且仅仅是这样的函数或运算, 可以同极限运算交换次序. 这就表明, 在数 a 上完成运算 f 所得到的数 $f(a)$, 可以依任意精确度借助于在量 a 的具有相应精确度的近似值 x 上完成运算 f 来逼近.

4° 如果注意到, 当 $a \in E$ 时, 点 a 的诸邻域 $U_E(a)$ 构成一个基 B_a(这与 a 是集合的极限点还是孤立点无关), 那么我们就会看到, 函数在点 a 处连续的定义 1, 与数 $f(a)$—— 函数在点 a 处的值 —— 是函数 f 关于这个基的极限的定义是一样的, 即

$$(f : E \to \mathbb{R} \text{在} a \in E \text{连续}) \Leftrightarrow (\lim_{B_a} f(x) = f(a)).$$

5° 但我们发现, 如果 $\lim\limits_{B_a} f(x)$ 存在, 那么, 由于对于任意的 $U_E(a)$ 都有 $a \in U_E(a)$, 于是这个极限非等于 $f(a)$ 不可.

这样一来, 函数 $f : E \to \mathbb{R}$ 在点 $a \in E$ 处连续等价于函数关于点 a 在 E 中的诸邻域 (但不是空心邻域) $U_E(a)$ 的基 B_a 有极限.

于是

$$(f : E \to \mathbb{R} \text{在} a \in E \text{连续}) \Leftrightarrow (\exists \lim_{B_a} f(x)).$$

6° 根据极限存在的柯西准则, 现在可以说, 函数在点 $a \in E$ 连续当且仅当对于任意的 $\varepsilon > 0$, 存在点 a 在 E 中的邻域 $U_E(a)$, 使在 $U_E(a)$ 上函数的振幅 $\omega(f, U_E(a))$ 小于 ε.

定义 2 量 $\omega(f, a) = \lim\limits_{\delta \to 0+0} \omega(f, U_E^\delta(a))$(其中 $U_E^\delta(a)$ 是点 a 在集 E 中的 δ 邻域) 叫做函数 $f : E \to \mathbb{R}$ 在点 $a \in E$ 的**振幅**.

符号 $\omega(f, X)$ 已被使用, 它表示函数在集 X 上的振幅. 但是我们永远也不考虑函数在一个点所成的集上的振幅 (这个振幅显然等于零); 因此, 符号 $\omega(f; a)$, 其中 a 是一个点, 永远表示我们刚刚用定义 2 引入的函数在一点处的振幅这样一个概念.

函数在子集合上的振幅不超过它在集合上的振幅, 因此, $\omega(f,U_E^\delta(a))$ 是 δ 的非增函数. 由于它是非负的, 那么, 或者它当 $\delta\to+0+0$ 时有有限的极限, 或者对于任意的 $\delta>0,\omega(f,U_E^\delta(a))=+\infty$. 在后一种情形时, 自然认为 $\omega(f,a)=+\infty$.

7° 使用在 6° 中叙述的定义 2, 现在可以做这样的概括: 函数在一点处连续当且仅当它在这点处的振幅等于零. 写作:

$$(f:E\to\mathbb{R}\text{在}a\in E\text{连续})\Leftrightarrow(\omega(f,a)=0).$$

定义 3　称函数 $f:E\to\mathbb{R}$ 在集合 E 上连续, 如果它在集合 E 的每点都连续.

我们约定把一切在集合 E 上连续的实函数的集合记作 $C(E;\mathbb{R})$ 或简记为 $C(E)$.

我们已经讨论了连续函数的概念, 现在我们来考察一些例子.

例 1　如果 $f:E\to\mathbb{R}$ 是常值函数, 则 $f\in C(E)$. 这个论断是明显的, 因为 $f(E)=c\in V(c)$, 不管 $c\in\mathbb{R}$ 的邻域 $V(c)$ 怎样它都成立.

例 2　函数 $f(x)=x$ 在 $\mathbb{R}$ 上连续.

实际上, 对于任意的点 $x_0\in\mathbb{R}$, 只要 $|x-x_0|<\delta=\varepsilon$, 就有 $|f(x)-f(x_0)|=|x-x_0|<\varepsilon$.

例 3　函数 $f(x)=\sin x$ 在 $\mathbb{R}$ 上连续.

事实上, 对于任意的点 $x_0\in\mathbb{R}$, 只要 $|x-x_0|<\delta=\varepsilon$, 我们就得到

$$\begin{aligned}|\sin x-\sin x_0|&=\left|2\cos\frac{x+x_0}{2}\sin\frac{x-x_0}{2}\right|\\&\leqslant 2\left|\sin\frac{x-x_0}{2}\right|\leqslant 2\left|\frac{x-x_0}{2}\right|\\&=|x-x_0|<\varepsilon.\end{aligned}$$

我们使用了在第三章 §2 第 2 段 d 款例 9b) 中证明了的不等式 $|\sin x|\leqslant|x|$.

例 4　函数 $f(x)=\cos x$ 在 $\mathbb{R}$ 上连续.

实际上, 像上例中一样, 对于任意的点 $x_0\in\mathbb{R}$, 只要 $|x-x_0|<\delta=\varepsilon$, 我们就有

$$\begin{aligned}|\cos x-\cos x_0|&=\left|-2\sin\frac{x+x_0}{2}\sin\frac{x-x_0}{2}\right|\\&\leqslant 2\left|\sin\frac{x-x_0}{2}\right|\leqslant|x-x_0|<\varepsilon.\end{aligned}$$

例 5　函数 $f(x)=a^x$ 在 $\mathbb{R}$ 上连续.

实际上, 根据指数函数的性质 3) (见第三章 §2 第 2 段 d 款例 10a)), 在任意点 $x_0\in\mathbb{R}$ 处, 有

$$\lim_{x\to x_0}a^x=a^{x_0},$$

现在我们知道, 这等价于函数 a^x 在点 x_0 连续.

例 6 函数 $f(x) = \log_a x$ 在定义域 $\mathbb{R}_+ = \{x \in \mathbb{R} | x > 0\}$ 的任意点 $x_0 \in \mathbb{R}_+$ 处连续.

事实上, 根据对数函数的性质 3) (见第二章 §2 第 2 段 d 款例 10b)), 在任意点 $x_0 \in \mathbb{R}_+$ 处, 有

$$\lim_{\mathbb{R}_+ \ni x \to x_0} \log_a x = \log_a x_0,$$

这等价于函数 $\log_a x$ 在点 x_0 处连续.

我们顺便对任给的 $\varepsilon > 0$ 来求点 x_0 的这样的邻域 $U_{\mathbb{R}_+}(x_0)$, 使得在任意的点 $x \in U_{\mathbb{R}_+}(x_0)$ 处有

$$|\log_a x - \log_a x_0| < \varepsilon.$$

这个不等式等价于关系式

$$-\varepsilon < \log_a \frac{x}{x_0} < \varepsilon.$$

为确定起见, 设 $a > 1$; 那么最后的关系式等价于条件

$$x_0 a^{-\varepsilon} < x < x_0 a^{\varepsilon}.$$

区间 $]x_0 a^{-\varepsilon}, x_0 a^{\varepsilon}[$ 就是点 x_0 的待求的邻域. 有益的是注意到这个邻域不但依赖于量 ε, 而且也依赖于点 x_0 本身. 这一点在例 1~4 中不曾出现.

例 7 任何数列 $f: \mathbb{N} \to \mathbb{R}$ 都是在自然数集 $\mathbb{N}$ 上连续的函数, 因为集 $\mathbb{N}$ 的每个点都是它的孤立点.

2. 间断点 为了更好地掌握连续这个概念, 我们要弄清楚函数在它的不连续的点的邻域里会是怎样的.

定义 4 如果函数 $f: E \to \mathbb{R}$ 在集 E 的某点不连续, 则此点叫做函数 f 的间断点.

只要做出 "函数 $f: E \to \mathbb{R}$ 在点 $a \in E$ 连续" 这一命题的否定, 我们就得到 a 是函数 f 的间断点的定义的下列写法:

$$(a \in E \text{是函数} f \text{的间断点}) := (\exists V(f(a))\ \forall U_E(a)\ \exists x \in U_E(a)\ (f(x) \notin V(f(a)))).$$

换言之, $a \in E$是函数 $f: E \to \mathbb{R}$ 的间断点, 如果函数在点 a 的值 $f(a)$ 有这样一个邻域 $V(f(a))$, 使得对点 a 在集合 E 中的任何一个邻域 $U_E(a)$, 都存在点 $x \in U_E(a)$, 它的像 $f(x)$ 不含在 $V(f(a))$ 中.

这个定义的 $\varepsilon - \delta$ 形式是:

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x \in E\ (|x - a| < \delta \wedge |f(x) - f(a)| > \varepsilon).$$

我们来看几个例子.

例 8 函数 $f(x)=\mathrm{sgn}x$ 在异于零的任意点 $a\in\mathbb{R}$ 的一个邻域内为常数, 从而它在这个邻域内连续. 而在零的任何邻域内, 它的振幅都等于 2, 因此, $\omega(\mathrm{sgn},0)=2$. 这就表示 0 是函数 $\mathrm{sgn}x$ 的间断点. 我们还看出, 函数在点 0 处有左极限 $\lim\limits_{x\to 0-0}\mathrm{sgn}x=-1$ 及右极限 $\lim\limits_{x\to 0+0}\mathrm{sgn}x=1$, 但是, 这两个极限彼此不等, 并且它们都与函数在点 0 处的值 $\mathrm{sgn}0=0$ 不等. 这直接证实了 0 是函数的间断点.

例 9 函数 $f(x)=|\mathrm{sgn}x|$ 当 $x\to 0$ 时有极限 $\lim\limits_{x\to 0}|\mathrm{sgn}x|=1$, 但是 $f(0)=|\mathrm{sgn}0|=0$, 因此

$$\lim_{x\to 0}f(x)\neq f(0),$$

于是 0 是函数的间断点.

但是我们会发现, 在所给的情况下, 只要改变函数在点 0 处的值, 并令它等于 1, 我们就得到一个在点 0 处连续的函数, 也就是说, 我们排除了间断.

定义 5 如果函数 $f:E\to\mathbb{R}$ 的间断点 $a\in E$ 是这样的: 存在一个在点 a 连续的函数 $\tilde{f}:E\to\mathbb{R}$ 使得 $f|_{E\backslash a}=\tilde{f}|_{E\backslash a}$, 那么 a 叫做函数 $f:E\to\mathbb{R}$ 的*可去间断点*.

这样一来, 可去间断点的特征是: 极限 $\lim\limits_{E\ni x\to a}f(x)=A$ 存在, 但是 $A\neq f(a)$, 从而, 只要置

$$\tilde{f}(x)=\begin{cases}f(x), & 当 x\in E, x\neq a,\\ A, & 当 x=a,\end{cases}$$

我们就得到一个在点 a 连续的函数 $\tilde{f}:E\to\mathbb{R}$.

例 10 函数 $f(x)=\begin{cases}\sin\dfrac{1}{x}, & 当 x\neq 0,\\ 0, & 当 x=0\end{cases}$ 在点 0 间断. 并且当 $x\to 0$ 时, 它没有极限, 这是因为在第三章 §2 第 1 段例 5 中曾经证明了极限 $\lim\limits_{x\to 0}\sin\dfrac{1}{x}$ 不存在. 函数 $\sin\dfrac{1}{x}$ 的图像画在图 12 中.

例 8、9 和 10 阐明了下面的术语.

定义 6 点 $a\in E$ 叫做函数 $f:E\to\mathbb{R}$ 的*第一类间断点*, 如果极限

$$\lim_{E\ni x\to a-0}f(x)=:f(a-0),\ \lim_{E\ni x\to a+0}f(x)=:f(a+0)$$

都存在[①], 但它们之中至少有一个与函数在点 a 的值 $f(a)$ 不同.

---

①如果 a 是间断点, 那么, a 是集 E 的极限点. 不过可能有这样的情况, 即集 E 在点 a 的某个邻域中的所有的点都在点 a 的同一侧, 在这种情况下只考虑定义中所述的极限当中的一个.

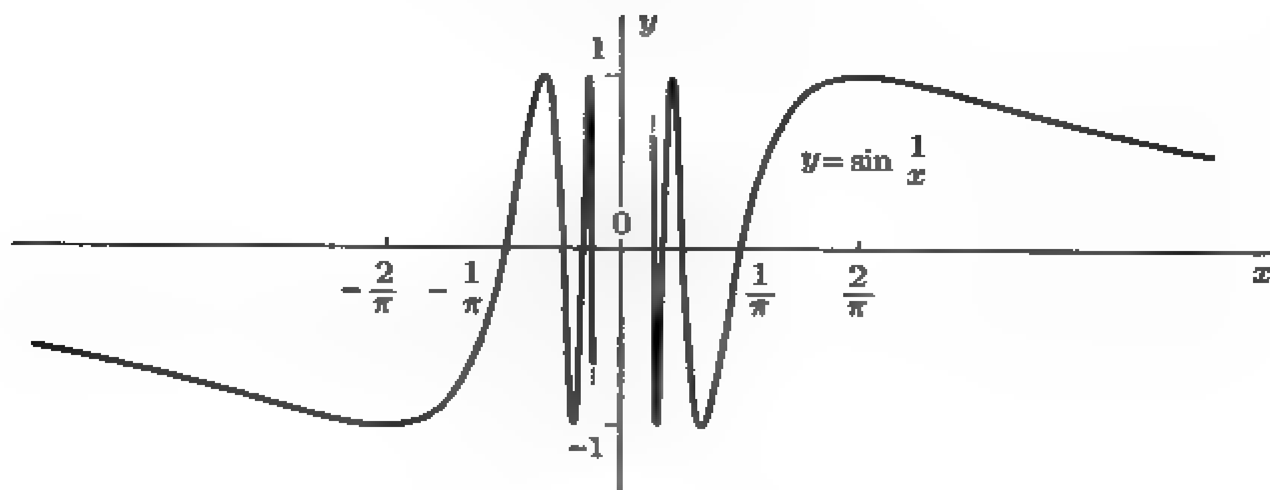


图 12

定义 7 如果 $a\in E$ 是函数 $f:E\to\mathbb{R}$ 的间断点, 并且在这点处定义 6 所述的两个极限至少有一个不存在, 那么 a 叫做第二类间断点.

这样一来, 任何一个间断点, 不是第一类间断点就是第二类间断点.

再举两个经典的例子.

例 11 函数

$$\mathcal{D}(x)=\begin{cases}1, 若 x\in\mathbb{Q},\\ 0, 若 x\in\mathbb{R}\backslash\mathbb{Q}\end{cases}$$

叫做狄利克雷[①]函数.

这个函数在所有的点都间断, 并且, 显然它的一切间断点都是第二类的, 因为在任何区间上都既有有理数也有无理数.

例 12 黎曼[②]函数

$$\mathcal{R}(x)=\begin{cases}\dfrac{1}{n}, 若 x=\dfrac{m}{n}\in\mathbb{Q}\left(\dfrac{m}{n}是既约分数\right),\\ 0, \ 若 x\in\mathbb{R}\backslash\mathbb{Q}.\end{cases}$$

我们指出, 对于无论怎样的点 $a\in\mathbb{R}$ 和它的有界邻域 $U(a)$ 以及无论怎样的数 $N\in\mathbb{N}$, 在 $U(a)$ 中仅有有限个有理数 $\dfrac{m}{n}, m\in\mathbb{Z}, n\in\mathbb{N}$, 使得 $n<N$.

因此, 只要缩小邻域, 就可以认为位于其中的一切有理数 (若 $a\in\mathbb{Q}$ 的话, 可能除去数 a) 的分母都比 N 大. 这样一来, 在任意点 $x\in\overset{\circ}{U}(a)$ 处, $|\mathcal{R}(x)|<\dfrac{1}{N}$.

这样我们就证明了, 在任意点 $a\in\mathbb{R}$ 处

$$\lim_{x\to a}\mathcal{R}(x)=0.$$

---

①狄利克雷 (P. G. Dirichlet)(1805—1859) —— 德国大分析家、数学家, 在高斯 (C. Gauss) 逝世后 (1855) 任哥廷根大学额内教授职.

②黎曼 (G. F. Riemann)(1826—1866) —— 卓越的德国数学家, 他的主要著作奠定了整个现代几何和现代分析领域的基础.

这表明, 黎曼函数在任意无理点处都连续. 在其他点处, 即在点 $x \in \mathbb{Q}$, 函数间断, 且所有这些点都是第一类间断点.

§2. 连续函数的性质

1. 局部性质 函数的那样一些性质叫做*局部的*, 它们是由函数在其定义域中点的任意小邻域中的性状确定的.

因此, 局部性质本身刻画的是当函数的自变量趋于所考察的点时, 函数在某种极限关系中的性状.

我们将指出连续函数的基本的局部性质.

定理 1 *设 $f: E \to \mathbb{R}$ 是在点 $a \in E$ 连续的函数, 那么下列论断成立:*

1° 函数 f 在点 a 的某个邻域 $U_E(a)$ 中有界.

2° 若 $f(a) \neq 0$, 则在点 a 的某个邻域 $U_E(a)$ 中函数的所有的值与 $f(a)$ 同时为正或同时为负.

3° 若函数 $g: U_E(a) \to \mathbb{R}$ 在点 a 的某个邻域中定义且和 $f: E \to \mathbb{R}$ 一样在点 a 处连续, 那么函数

a) $(f+g)(x) := f(x) + g(x)$,

b) $(f \cdot g)(x) := f(x) \cdot g(x)$,

c) $\left(\frac{f}{g}\right)(x) := \frac{f(x)}{g(x)}$ *(在 $g(a) \neq 0$ 的条件下)*

在点 a 的某个邻域中定义且在点 a 连续.

4° 若函数 $g: Y \to \mathbb{R}$ 在点 $b \in Y$ 连续, 而函数 $f: E \to Y$ 有 $f(a) = b$ 且 f 在点 a 连续, 那么, 复合函数 $(g \circ f)$ 在 E 上定义且在点 a 连续.

◀ 为证定理只需想到 (见 §1), 函数 f 或 g 在定义域的某点 a 连续等价于此函数关于点 a 的邻域基 $\mathcal{B}_a$ 的极限存在且等于函数在该点 a 处的值:

$$\lim_{\mathcal{B}_a} f(x) = f(a), \quad \lim_{\mathcal{B}_a} g(x) = g(a).$$

于是, 定理 1 的论断 1°、2°、3° 直接从函数在一点连续的定义和函数极限的相应性质推出.

只是需要说明, 比式 $\frac{f(x)}{g(x)}$ 确实在点 a 的某个邻域 $\tilde{U}_E(a)$ 中有定义. 这是因为按条件 $g(a) \neq 0$ 且根据定理的论断 2°, 存在一个邻域 $\tilde{U}_E(a)$, 在它的任意一点都有 $g(x) \neq 0$, 即 $\frac{f(x)}{g(x)}$ 在 $\tilde{U}_E(a)$ 中有定义.

定理 1 的论断 4° 是关于复合函数的极限的定理的推论. 根据关于复合函数的极限的定理

$$\lim_{\mathcal{B}_a} (g \circ f)(x) = \lim_{\mathcal{B}_b} g(y) = g(b) = g(f(a)) = (g \circ f)(a),$$

这等价于 $(g \circ f)$ 在点 a 连续.

不过, 要使用关于复合函数的极限的定理, 必须验证对于基 $\mathcal{B}_b$ 的任何一个元 $U_Y(b)$ 都存在基 $\mathcal{B}_a$ 的这样的元 $U_E(a)$, 使得 $f(U_E(a)) \subset U_Y(b)$. 事实上, 若 $U_Y(b) = Y \cap U(b)$, 则按函数 $f: E \to Y$ 在点 a 连续的定义, 对于邻域 $U(b) = U(f(a))$, 存在点 a 在集 E 中的邻域 $U_E(a)$, 使得 $f(U_E(a)) \subset U(f(a))$. 由于 f 是从 E 映到 Y 中的, 那么, $f(U_E(a)) \subset Y \cap U(f(a)) = U_Y(b)$, 于是我们证实了使用关于复合函数的极限的定理的合理性. ►

例 1 代数多项式 $P(x) = a_0x^n + a_1x^{n-1} + \cdots + a_n$ 是在 $\mathbb{R}$ 上连续的函数.

实际上, 从定理 1 的 3° 按归纳法可以推出有限个在某点连续的函数的和和积都是在这点连续的函数. 在 §1 的例 1 和例 2 中我们已经验证了常数函数及函数 $f(x) = x$ 在 $\mathbb{R}$ 上连续. 那么函数 $a \cdot x^m = a \cdot \underbrace{x \cdot \cdots \cdot x}_{m\text{个}}$ 也在 $\mathbb{R}$ 上连续, 因此多项式 $P(x)$ 在 $\mathbb{R}$ 上连续.

例 2 有理函数 $R(x) = \dfrac{P(x)}{Q(x)}$—— 多项式的比, 在它有定义即 $Q(x) \neq 0$ 的地方处处连续. 这是从例 1 和定理 1 的论断 3° 推出的.

例 3 有限个连续函数的复合, 在自己的定义域的任何一点都是连续的. 这是从定理 1 的论断 4° 用归纳法推出的. 特别地, 例如函数 $e^{\sin^2(\ln|\cos x|)}$ 在 $\mathbb{R}$ 上除去它没有定义的那些点即点 $\frac{\pi}{2}(2k+1), k \in \mathbb{Z}$, 之外是处处连续的.

2. 连续函数的整体性质 用描述性的语言来说, 与函数的整个定义域有关的性质叫做函数的整体性质.

定理 2 (波尔察诺–柯西中间值定理) *若函数在闭区间上连续, 在它的两端点取异号的值, 则在区间上有一点使函数在该点的值为零.*

用逻辑符号此定理有如下写法①:

$$(f \in C[a, b]) \wedge (f(a) \cdot f(b) < 0) \Rightarrow (\exists c \in [a, b]\, f(c) = 0).$$

◄ 把区间 $[a, b]$ 平分成两半. 若在分点处函数不等于零, 则平分所得的两个区间中有一个, 在它的两端函数取异号的值. 现在像处理原来的区间 $[a, b]$ 一样地处理这个区间, 即把它平分成两半, 并且进一步继续这个过程.

那么, 我们或者在某一步得到一个点 $c \in [a, b]$, 使得 $f(c) = 0$, 或者得到一个闭区间套 $\{I_n\}$, 这些区间的长趋于零. 在后一种情况下, 根据关于闭区间套的引理, 存在唯一的一个点 $c \in [a, b]$, 它是所有这些闭区间的公共点. 按闭区间套的结构, 存在

①我们记得, 符号 $C(E)$ 表示一切在集 E 上连续的函数的集合. 在 $E = [a, b]$ 的情形下, 常代替 $C([a, b])$ 而简写 $C[a, b]$.

由诸区间 I_n 的端点组成的两个数列 $\{x'_n\}$ 和 $\{x''_n\}$, 使得 $f(x'_n) < 0, f(x''_n) > 0$,

$$\lim_{n\to\infty} x'_n = \lim_{n\to\infty} x''_n = c.$$

根据极限的性质和连续的定义, 我们得到

$$\lim_{n\to\infty} f(x'_n) = f(c) \leqslant 0, \lim_{n\to\infty} f(x''_n) = f(c) \geqslant 0.$$

于是 $f(c) = 0$. ►

定理 2 的注 1° 定理的证明提供了一个求方程 $f(x) = 0$ 在这样的闭区间上的根的最简单的算法, 在这个闭区间的两端连续函数 f 取异号的值.

2° 于是定理 2 断定, 函数在连续变化过程中, 不可能一直不取零值且从正值变到负值或从负值变到正值.

3° 应该审慎地对待 2° 中那样的文字表述, 因为这种表述中所含的意思常常比要表达的更多. 例如, 我们来考察在区间 $[0, 1]$ 上等于 -1 并在区间 $[2, 3]$ 上等于 1 的函数. 显然, 这个函数在自己的定义域上连续且取不同符号的值, 但处处不为零. 这个注解表明, 定理 2 所表达的连续函数的性质实际上产生于函数定义域的某种性质 (以后会弄明白, 这种性质是: 这个集合应该是连通的).

定理 2 的推论 *若函数 φ 在区间上连续且在区间的点 a 和 b 取值 $\varphi(a) = A$ 和 $\varphi(b) = B$, 则对于介于 A 和 B 之间的任意数 C, 存在介于 a 和 b 之间的点 c, 使得 $\varphi(c) = C$.*

◄ 以 a, b 为端点的闭区间 I 含在我们的区间中, 因此函数 $f(x) = \varphi(x) - C$ 在 I 上有定义、连续且由于

$$f(a) \cdot f(b) = (A - C)(B - C) < 0,$$

根据定理 2, 在 a 和 b 之间有点 c 使

$$f(c) = \varphi(c) - C = 0. \blacktriangleright$$

定理 3 (魏尔斯特拉斯最大值定理) *在闭区间上连续的函数在此区间上有界. 同时, 在区间上存在使函数取得最大值的点, 也存在使函数取得最小值的点.*

◄ 设 $f : E \to \mathbb{R}$ 是闭区间 $E = [a, b]$ 上的连续函数. 根据连续函数的局部性质 (见定理 1), 对于任意的点 $x \in E$, 存在这样的邻域 $U(x)$, 使得在集 $U_E(x) = E \cap U(x)$ 上函数有界. 对于一切点 $x \in E$ 所构造的这样的邻域 $U(x)$ 的全体组成闭区间 $[a, b]$ 的开覆盖, 根据有限覆盖引理可以从这个开覆盖中取出有限个开区间 $U(x_1)$、$\cdots$、$U(x_n)$, 它们覆盖闭区间 $[a, b]$. 由于在集 $E \cap U(x_k) = U_E(x_k)$ 上函数有界, 即 $m_k \leqslant f(x) \leqslant M_k$, 其中 m_k、$M_k \in \mathbb{R}$ 且 $x \in U_E(x_k)$, 那么, 在任意点 $x \in E = [a, b]$ 处, 有

$$\min\{m_1, \cdots, m_k\} \leqslant f(x) \leqslant \max\{M_1, \cdots, M_k\}.$$

函数在闭区间 $[a,b]$ 上的有界性被证实了.

现设 $M=\sup\limits_{x\in E} f(x)$. 假定在任意点 $x\in E(f(x)<M)$. 那么, E 上的连续函数 $M-f(x)$ 尽管 (根据 M 的定义) 可以任意接近于零却处处不为零. 于是函数 $\dfrac{1}{M-f(x)}$ 一方面根据连续函数的局部性质是在 E 上连续的, 而另一方面却是在 E 上无界的, 这与已经证明了的闭区间上连续函数的有界性矛盾.

于是存在点 $x_M\in[a,b]$, 使 $f(x_M)=M$.

类似地, 只要考虑 $m=\inf\limits_{x\in E} f(x)$ 和辅助函数 $\dfrac{1}{f(x)-m}$ 就可以证明存在点 $x_m\in[a,b]$, 在这点处 $f(x_m)=m$. ▶

我们注意到, 例如 $f_1(x)=x$、$f_2(x)=\dfrac{1}{x}$ 都在开区间 $E=]0,1[$ 连续, 但 f_1 在 E 上既无最大值亦无最小值, 而函数 f_2 在 E 上无界. 因此, 定理 2 所说的连续函数的性质也和定义域的某种性质有关, 这种性质就是: 从由集合的点的邻域构成的它的覆盖中可以取出有限覆盖. 以后我们称这样的集合为紧集.

在转入下一个定理之前我们给出

定义 1 函数 $f:E\to\mathbb{R}$ 叫做是在集 $E\subset\mathbb{R}$ 上一致连续的, 如果对于任意的数 $\varepsilon>0$, 存在数 $\delta>0$, 使得对于任意满足 $|x_1-x_2|<\delta$ 的点 x_1、$x_2\in E$, $|f(x_1)-f(x_2)|<\varepsilon$ 成立.

简言之

$$(f:E\to\mathbb{R}\text{一致连续}):=$$
$$=(\forall\varepsilon>0\ \exists\delta>0\ \forall x_1\in E\ \forall x_2\in E\ (|x_1-x_2|<\delta\Rightarrow|f(x_1)-f(x_2)|<\varepsilon)).$$

我们来讨论一致连续函数的概念.

1° 若函数在集上一致连续, 则它在集的任意点处连续. 实际上, 只要在所述定义中置 $x_1=x$ 和 $x_2=a$, 我们就见到函数 $f:E\to\mathbb{R}$ 在点 $a\in E$ 连续的定义已满足.

2° 一般说来函数的连续性并不蕴含一致连续性.

例 4 我们已不止一次地遇到的开区间 $]0,1[=E$ 上的函数 $f(x)=\sin\dfrac{1}{x}$ 是连续的. 但是, 在点 0 于集 E 的任何邻域之中, 函数既取值 -1 也取值 1, 因此, 当 $\varepsilon<2$ 时, 它已不满足条件 $|f(x_1)-f(x_2)|<\varepsilon$.

因此, 写出函数不一致连续的性质的明显形式是有益的:

$$(f:E\to\mathbb{R}\text{不是一致连续的}):=$$
$$=(\exists\varepsilon>0\ \forall\delta>0\ \exists x_1\in E\ \exists x_2\in E\ (|x_1-x_2|<\delta\wedge|f(x_1)-f(x_2)|\geqslant\varepsilon)).$$

所考察的例子显示了函数在集合上的连续性和一致连续性这二者之间的明显差异. 为了表明在一致连续的定义中是什么地方引起了这个差异的, 我们给出表示函

数 $f: E \to \mathbb{R}$ 在集 E 上连续的详细的写法:

$$(f: E \to \mathbb{R}\text{在}E\text{上连续}) := (\forall a \in E\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in E$$
$$(|x-a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon)).$$

这样一来, 这里的 δ 是按点 $a \in E$ 和数 ε 来选择的. 因此对于固定的 ε 它可能随着点的改变而改变, 这正像在例 1 中所考察的函数 $\sin\frac{1}{x}$ 的情形下所发生的那样, 或者也如在函数的整个定义域中来考察 $\log_a x$ 或 a^x 的情形那样.

而在一致连续的情况下, 保证可以只按 $\varepsilon > 0$ 来选择 δ, 使得在一切点 $a \in E$ 处, 从当 $x \in E$ 满足 $|x-a| < \delta$ 必定得到 $|f(x) - f(a)| < \varepsilon$.

例 5 若函数 $f: E \to \mathbb{R}$ 在固定点 $x_0 \in \mathbb{R}$ 的任意邻域内都无界, 则它不是一致连续函数.

实际上, 那时对于任意的 $\delta > 0$, 在 x_0 的 $\frac{\delta}{2}$-邻域内找得到点 x_1、$x_2 \in E$, 尽管 $|x_1 - x_2| < \delta$, 都有 $|f(x_1) - f(x_2)| > 1$.

在集 $\mathbb{R}\backslash 0$ 上考察函数 $f(x) = \frac{1}{x}$, 情况就是这样. 在这种情况下, $x_0 = 0$.

对于定义在正数集上且在点 $x_0 = 0$ 的邻域内无界的函数 $\log_a x$, 情况也是这样.

例 6 在 $\mathbb{R}$ 上连续的函数 $f(x) = x^2$, 在 $\mathbb{R}$ 上不是一致连续.

事实上, 在点 $x'_n = \sqrt{n+1}, x''_n = \sqrt{n}$ 处有

$$f(x'_n) = n+1, \quad f(x''_n) = n,$$

其中 $n \in \mathbb{N}$, 因此 $f(x'_n) - f(x''_n) = 1$. 然而

$$\lim_{n\to\infty}(\sqrt{n+1} - \sqrt{n}) = \lim_{n\to\infty}\frac{1}{\sqrt{n+1}+\sqrt{n}} = 0,$$

因此, 对于任意的 $\delta > 0$, 存在点 x'_n、x''_n, 使得 $|x'_n - x''_n| < \delta$, 同时 $f(x'_n) - f(x''_n) = 1$.

例 7 在 $\mathbb{R}$ 上连续且有界的函数 $f(x) = \sin x^2$ 在 $\mathbb{R}$ 上不一致连续, 因为在点 $x'_n = \sqrt{\frac{\pi}{2}(n+1)}$、$x''_n = \sqrt{\frac{\pi}{2}n}$ 处有

$$|f(x'_n) - f(x''_n)| = 1,$$

其中 $n \in \mathbb{N}$, 同时

$$\lim_{n\to\infty}|x'_n - x''_n| = 0.$$

在对于函数一致连续的概念及连续与一致连续的对比做了这些讨论之后, 我们现在就可以认识到下述定理的价值.

定理 4 (康托尔海涅关于一致连续的定理) *在闭区间上连续的函数在此区间上一致连续.*

◀ 设 $f: E \to \mathbb{R}$ 是给定的函数; $E = [a, b]$ 且 $f \in C(E)$. 由于 f 在任意点 $x \in E$ 处连续, 所以 (见 §1 第 1 段 6°) 对于 $\varepsilon > 0$ 可以找到 x 点的这样的邻域 $U^\delta(x)$, 使得函数 f 在集合 $U_E^\delta(x) = E \cap U^\delta(x)$ 上的振幅 $\omega(f, U_E^\delta(x))$ 小于 ε, 此处集 $U_E^\delta(x)$ 是函数的定义域中那些在 $U^\delta(x)$ 中的点所成的集合. 对于每点 $x \in E$ 都构造具有这种性质的邻域. 这里量 δ 可以随着点的改变而改变, 因此用符号 $U^{\delta(x)}(x)$ 表示所构造的邻域更正确些, 虽然比较累赘. 不过, 由于整个符号都由点 x 决定, 可以约定使用下列简化的记法:

$$U(x) = U^{\delta(x)}(x) \text{ 和 } V(x) = U^{\frac{1}{2}\delta(x)}(x).$$

诸开区间 $V(x), x \in E$, 其全体构成闭区间 $E = [a, b]$ 的覆盖, 从这个覆盖中, 按有限覆盖引理可以取出有限覆盖 $V(x_1), \cdots, V(x_n)$. 设

$$\delta = \min\left\{\frac{1}{2}\delta(x_1), \cdots, \frac{1}{2}\delta(x_n)\right\}.$$

我们来证明, 对于任意的点 x'、$x'' \in E$, 只要 $|x' - x''| < \delta$ 就成立 $|f(x') - f(x'')| < \varepsilon$. 实际上, 由于开区间组 $V(x_1), \cdots, V(x_n)$ 覆盖 E, 必存在这个组的一个区间 $V(x_i)$ 含有点 x', 即 $|x' - x_i| < \dfrac{1}{2}\delta(x_i)$. 但在这种情况下,

$$\begin{aligned} |x'' - x_i| &\leqslant |x' - x''| + |x' - x_i| < \delta + \frac{1}{2}\delta(x_i) \\ &\leqslant \frac{1}{2}\delta(x_i) + \frac{1}{2}\delta(x_i) = \delta(x_i). \end{aligned}$$

因此, x'、$x'' \in U_E^{\delta(x_i)}(x_i) = E \cap U^{\delta(x_i)}(x_i)$, 从而 $|f(x') - f(x'')| \leqslant \omega(f, U_E^{\delta(x_i)}(x_i)) < \varepsilon$. ▶

前面所举的例子表明, 康托尔定理实质上依赖于函数的定义域的某种性质. 从证明见到, 正像在定理 3 中一样, 这种性质是: 从集 E 的任何由它的点的邻域所成的覆盖中都可以取出有限覆盖.

现在, 当定理 4 已证明之后, 重新回到上面已分析过的连续但不一致连续的函数的例子是有益的, 将能弄清楚, 譬如函数 $\sin x^2$ 是如何依照康托尔定理在实轴的每个闭区间上都一致连续而在 $\mathbb{R}$ 上却不一致连续. 此事之原因与连续函数一般说来不必一致连续的原因完全类似. 这一次我们让读者自己弄清楚这个问题.

现在我们转向本节的最后一个定理 —— 关于反函数的定理. 我们面临的问题是弄清楚, 在怎样的条件下在闭区间上连续的实值函数有反函数, 以及在什么情况下这个反函数连续.

命题 1 *闭区间 $E = [a, b]$ 到 $\mathbb{R}$ 中的连续映射 $f: E \to \mathbb{R}$ 是单射的充要条件是函数 f 在闭区间 $[a, b]$ 上严格单调.*

◀ 若函数 f 在任意的集合 $E \subset \mathbb{R}$ 中是增的或减的, 则映射 $f: E \to \mathbb{R}$ 显然是单射: 在集 E 的不同的点处函数取不同的值.

所以，命题 1 的最本质的内容在于说，任何闭区间的连续单射 $f:[a,b]\to\mathbb{R}$ 都是严格单调的函数.

假设不是这样，我们就找得到区间 $[a,b]$ 的三个点 $x_1<x_2<x_3$，使 $f(x_2)$ 不在 $f(x_1)$ 和 $f(x_3)$ 之间. 在这种情况下，或者 $f(x_3)$ 在 $f(x_1)$ 和 $f(x_2)$ 之间，或者 $f(x_1)$ 在 $f(x_2)$ 和 $f(x_3)$ 之间. 为确定起见，设所述两种可能情形的后一种情形发生. 按条件函数 f 在闭区间 $[x_2,x_3]$ 上连续，因此（见定理 2 的推论）在这个区间上存在点 x_1' 使 $f(x_1')=f(x_1)$. 这样一来，$x_1<x_1'$ 且 $f(x_1)=f(x_1')$，这与映射的单射性质不相容. $f(x_3)$ 在 $f(x_1)$ 和 $f(x_2)$ 之间的情形可类似地分析. ►

命题 2 *每个定义在数集 $X\subset\mathbb{R}$ 上的严格单调函数 $f:X\to\mathbb{R}$ 都有反函数 $f^{-1}:Y\to\mathbb{R}$，它定义在函数 f 的值的集合 $Y=f(X)$ 上且在 Y 上具有与函数 f 在集 X 上相同的单调性.*

◄ 映射 $f:X\to Y=f(X)$ 是满射，即是到集合 Y 上的映射，为确定起见设 $f:X\to Y$ 在 X 上增. 在这种情况下

$$\forall x_1\in X\quad \forall x_2\in X(x_1<x_2\Leftrightarrow f(x_1)<f(x_2)). \tag{1}$$

这样一来，映射 $f:X\to Y$ 在不同的点取不同的值，即它是单射. 结果，$f:X\to Y$ 是满单射，即 f 是 X 到 Y 的双方单值映射. 这就是说，当 $y=f(x)$ 时，$x=f^{-1}(y)$ 这一公式给出的逆映射 $f^{-1}:Y\to X$ 有定义.

把映射 $f^{-1}:Y\to X$ 的定义和关系式 (1) 相对照，我们得到关系式

$$\forall y_1\in Y\quad \forall y_2\in Y\ (f^{-1}(y_1)<f^{-1}(y_2)\Leftrightarrow y_1<y_2), \tag{2}$$

它表明函数 f^{-1} 在其定义域上是增的.

$f:X\to Y$ 在 X 上减的情形显然可类似地分析. ►

由已证明的命题 2 可见，如果想知道实值函数的反函数的连续性，研究单调函数连续的条件是有用的.

命题 3 *在集 $E\subset\mathbb{R}$ 上单调的函数 $f:E\to\mathbb{R}$ 只可能有第一类间断点.*

◄ 为确定起见，设 f 是不减函数. 假设 $a\in E$ 是 f 的间断点. 由于 a 不能是 E 的孤立点，所以 a 至少是集

$$E_a^-=\{x\in E|x<a\}\text{ 和 }E_a^+=\{x\in E|a<x\}$$

之中一个的极限点. 由于 f 是不减函数，所以对任意点 $x\in E_a^-$ 都有 $f(x)\leqslant f(a)$，于是 f 在 E_a^- 上的限制 $f|_{E_a^-}$ 是不减的有上界的函数. 那么，存在极限

$$\lim_{E_a^-\ni x\to a}(f|_{E_a^-})(x)=\lim_{E\ni x\to a-0}f(x)=f(a-0).$$

类似地可以证明存在极限

$$\lim_{E\ni x\to a+0} f(x) = f(a+0),$$

如果 a 是集 E_a^+ 的极限点的话.

当 f 是非增函数时, 可以重述所进行的证明, 也可以转而考虑函数 $-f$ 而把问题归结到已经考虑过的情形. ►

推论 1 *若点 a 是单调函数 $f: E\to\mathbb{R}$ 的间断点, 则极限*

$$\lim_{E\ni x\to a-0} f(x) = f(a-0), \quad \lim_{E\ni x\to a+0} f(x) = f(a+0)$$

之中至少有一个是确定的; 如果 f 是非减的 (或者非增的) 那么不等式

$$f(a-0)\leqslant f(a)\leqslant f(a+0)(\text{或者} f(a-0)\geqslant f(a)\geqslant f(a+0))$$

之中至少有一个成立严格的不等号; 在这个严格的不等式所确定的开区间中不含有函数的值; 对应于单调函数的不同的间断点的这种开区间彼此不相交.

◄ 实际上, 若 a 是间断点, 则它是集 E 的极限点且根据命题 3 它是第一类间断点. 因此, 基

$$E\ni x\to a-0 \quad \text{和} \quad E\ni x\to a+0$$

中至少有一个是有定义的且沿着这个基 (当两个基都有定义时, 则分别沿着每个基) 函数 f 有极限. 为确定起见, 设 f 是非减函数. 由于 a 是间断点, 那么, 不等式 $f(a-0)\leqslant f(a)\leqslant f(a+0)$ 中至少有一个实际上是严格的不等式. 由于当 $x\in E$ 且 $x<a$ 时,

$$f(x)\leqslant \lim_{E\ni x\to a-0} f(x) = f(a-0),$$

而类似地当 $x\in E$ 且 $a<x$ 时, $f(a+0)\leqslant f(x)$, 所以, 由严格不等式 $f(a-0)<f(a)$ 或 $f(a)<f(a+0)$ 确定的开区间实际上不含有函数的值. 设 a_1、a_2 是函数的两个不同的间断点, 并设 $a_1<a_2$, 那么由于函数 f 不减, 我们有

$$f(a_1-0)\leqslant f(a_1)\leqslant f(a_1+0)\leqslant f(a_2-0)\leqslant f(a_2)\leqslant f(a_2+0).$$

由此推出, 相应于不同的间断点的那些不含函数值的开区间是不相交的. ►

推论 2 *单调函数的间断点的集合至多可数.*

◄ 根据推论 1, 我们把单调函数的每个间断点和一个这样的开区间联系起来, 这个开区间是由函数在间断点处的值和函数当自变量从右边或左边趋于间断点时的一个极限来确定的. 这些开区间不相交. 但在直线上不相交的开区间的集合不可能比可数集大. 事实上, 在这些开区间的每个当中可取一个有理点, 那时, 开区间的集合

就与全体有理数所成的可数集 $\mathbb{Q}$ 的子集同势. 这就是说, 它本身至多可数. 单调函数的间断点的集合按结构是与这个开区间的集合同势的, 因此和它一样是至多可数集. ►

命题 4 (单调函数的连续性准则) *在闭区间 $E=[a,b]$ 上的单调函数 $f:E\to\mathbb{R}$ 在 E 上连续当且仅当它的值的集合 $f(E)$ 本身是一个以 $f(a)$ 和 $f(b)$ 为端点*[①]*的闭区间.*

◄ 若 f 是连续的单调函数, 则根据 f 的单调性, 函数在闭区间 $[a,b]$ 中所取的一切值都介于它在区间的端点所取的值 $f(a)$ 和 $f(b)$ 之间. 根据函数的连续性, 它必定也取到一切介于 $f(a)$ 和 $f(b)$ 之间的中间值. 这样一来, 在闭区间 $[a,b]$ 上单调且连续的函数的值的集合实际上是以 $f(a)$ 和 $f(b)$ 为端点的闭区间.

现在我们来证明逆命题. 设 f 是在闭区间 $[a,b]$ 上单调的函数. 若它在某点 $c\in[a,b]$ 处间断, 则据命题 3 的推论 1, 开区间

$$]f(c-0),f(c)[,\quad]f(c),f(c+0)[$$

中的一个是确有定义的, 且它不含有我们的函数的值, 但据函数的单调性, 这个开区间是包含在以 $f(a)$ 和 $f(b)$ 为端点的闭区间之中的. 因此, 若在闭区间 $[a,b]$ 上单调的函数具有哪怕只是一个间断点, 则以 $f(a)$、$f(b)$ 为端点的闭区间就不能整个地含在函数的值域之中. ►

定理 5 (关于反函数) *集 $X\subset\mathbb{R}$ 上严格单调的函数 $f:X\to\mathbb{R}$ 有反函数 $f^{-1}:Y\to\mathbb{R}$, 它定义在函数 f 的值的集合 $Y=f(X)$ 上. 函数 $f^{-1}:Y\to\mathbb{R}$ 是单调的, 而且它在 Y 上的单调形式和函数 $f:X\to\mathbb{R}$ 在集 X 上的单调形式相同.*

此外, 若 X 是闭区间 $[a,b]$, 且函数在此集上连续, 则集 $Y=f(X)$ 是以 $f(a)$ 和 $f(b)$ 为端点的闭区间, 且函数 $f^{-1}:Y\to\mathbb{R}$ 在集 Y 上连续.

◄ 定理中关于当 $X=[a,b]$ 且 f 连续时, $Y=f(X)$ 是以 $f(a)$、$f(b)$ 为端点的闭区间的结论, 从上面已证的命题 4 推出. 剩下的是验证 $f^{-1}:Y\to\mathbb{R}$ 是连续函数. 但是 f^{-1} 在 Y 上单调, Y 是闭区间, 且 $f^{-1}(Y)=X=[a,b]$ 也是闭区间. 根据命题 4 我们断定函数 f^{-1} 在以 $f(a)$、$f(b)$ 为端点的闭区间 Y 上连续. ►

例 8 函数 $y=f(x)=\sin x$ 在闭区间 $\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$ 上递增且连续. 这就意味着, 这个函数在闭区间 $\left[-\frac{\pi}{2},\frac{\pi}{2}\right]$ 上的限制具有反函数 $x=f^{-1}(y)$, 记为 $x=\arcsin y$, 它定义在闭区间 $\left[\sin\left(-\frac{\pi}{2}\right),\sin\frac{\pi}{2}\right]=[-1,1]$ 上, 由 $-\frac{\pi}{2}$ 递增到 $\frac{\pi}{2}$, 且在这个闭区间上连续.

①这里, 若 f 是非减的, 则 $f(a)\leqslant f(b)$; 若 f 是非增的, 则 $f(b)\leqslant f(a)$.

例 9 类似地, 函数 $y = \cos x$ 在闭区间 $[0,\pi]$ 上的限制是递减连续函数. 根据定理 5 它有反函数, 记为 $x = \arccos y$, 这个函数定义在闭区间 $[-1,1]$ 上, 且在这个区间上由值 π 递减到值 0.

例 10 函数 $y = \mathrm{tg} x$ 在开区间 $X = \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ 上的限制是由 $-\infty$ 向 $+\infty$ 递增的连续函数, 根据定理 5 的第一部分, 它有定义在整个数轴 $Y = \mathbb{R}$ 上的反函数, 记为 $x = \mathrm{arctg}\, y$, 且在由它自己的值所成的开区间 $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ 的范围内递增. 为了证明函数 $x = \mathrm{arctg}\, y$ 在它的定义域的任意一点 y_0 处连续, 我们取点 $x_0 = \mathrm{arctg}\, y_0$ 和开区间 $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ 内包含点 x_0 的闭区间 $[x_0 - \varepsilon, x_0 + \varepsilon]$. 如果

$$x_0 - \varepsilon = \mathrm{arctg}(y_0 - \delta_1)$$

以及

$$x_0 + \varepsilon = \mathrm{arctg}(y_0 + \delta_2),$$

那么, 根据函数 $x = \mathrm{arctg}\, y$ 的递增性质可以断定, 当 $y \in \mathbb{R}$ 使得 $y_0 - \delta_1 < y < y_0 + \delta_2$ 时, 将有

$$x_0 - \varepsilon < \mathrm{arctg}\, y < x_0 + \varepsilon.$$

于是当 $-\delta_1 < y - y_0 < \delta_2$ 时,

$$|\mathrm{arctg}\, y - \mathrm{arctg}\, y_0| < \varepsilon,$$

而当 $|y - y_0| < \delta = \min\{\delta_1, \delta_2\}$ 时, 当然更有此式成立. 这就验证了函数 $x = \mathrm{arctg}\, y$ 在点 $y_0 \in \mathbb{R}$ 处的连续性.

例 11 通过类似于上例中所进行的讨论, 我们可以明白, 由于函数 $y = \mathrm{ctg}\, x$ 在开区间 $]0,\pi[$ 上的限制是由 $+\infty$ 递减到 $-\infty$ 的连续函数, 所以它有反函数, 记为 $x = \mathrm{arcctg}\, y$, 此函数定义在全数轴 $\mathbb{R}$ 上, 在由它自己的值所成的开区间 $]0,\pi[$ 的范围内从 π 到 0 递减, 且在 $\mathbb{R}$ 上连续.

注 在作互反函数 $y = f(x)$ 和 $x = f^{-1}(y)$ 的图像时, 注意到以下事实是有益的: 平面上在同一个坐标系中 (这个坐标系中只标出第一和第二坐标轴, 而不标 x 轴或 y 轴) 坐标为 $(x, f(x)) = (x, y)$ 和 $(y, f^{-1}(y)) = (y, x)$ 的两点关于第一象限的平分线是对称的.

因此, 互反函数在同一个坐标系中的图像关于这条平分线是对称的.

练 习

1. 证明

   a) 若 $f \in C(A)$ 且 $B \subset A$, 则 $f|_B \in C(B)$.

b) 若函数 $f: E_1 \cup E_2 \to \mathbb{R}$ 使 $f|_{E_i} \in C(E_i)(i=1,2)$, 则并不总有 $f \in C(E_1 \cup E_2)$.

c) 黎曼函数 $\mathcal{R}$ 和它在有理数集上的限制 $\mathcal{R}|_{\mathbb{Q}}$ 一样, 在集 $\mathbb{Q}$ 的每点, 除零外, 都间断, 并且所有的间断点都是可去间断点 (见 §1 例 12).

2. 证明, 若函数 $f \in C[a,b]$, 则函数

$$m(x) = \min_{a \leqslant t \leqslant x} f(t), \quad M(x) = \max_{a \leqslant t \leqslant x} f(x)$$

也在闭区间 $[a,b]$ 上连续.

3. a) 证明, 开区间上的单调函数的反函数在自己的定义域上连续.

b) 构造一个具有可数个间断点的单调函数.

c) 证明, 若函数 $f: X \to Y$ 和 $f^{-1}: Y \to X$ 互反 (这里 X、Y 都是 $\mathbb{R}$ 的子集) 且 f 在点 $x_0 \in X$ 连续, 那么, 由此还不能推出函数 f^{-1} 在点 $y_0 = f(x_0) \in Y$ 连续.

4. 证明

a) 若 $f \in C[a,b]$ 且 $g \in C[a,b]$, 同时 $f(a) < g(a)$ 且 $f(b) > g(b)$, 那么, 存在点 $c \in [a,b]$, 在这点处 $f(c) = g(c)$.

b) 任何闭区间到自身的连续映射 $f: [0,1] \to [0,1]$ 都有不动点, 即点 $x \in [0,1]$ 使 $f(x) = x$.

c) 若从闭区间到自身的连续映射 f 和 g 是可交换的, 即 $f \circ g = g \circ f$, 则它们有共同的不动点.

d) 连续映射 $f: \mathbb{R} \to \mathbb{R}$ 可以没有不动点.

e) 连续映射 $f:]0,1[\to]0,1[$ 可以没有不动点.

f) 若映射 $f: [0,1] \to [0,1]$ 连续, $f(0) = 0, f(1) = 1$ 且在 $[0,1]$ 上 $(f \circ f)(x) \equiv x$, 则 $f(x) \equiv x$.

5. 证明任意一个在闭区间上连续的函数的值的集合也是闭区间.

6. 证明

a) 若映射 $f: [0,1] \to [0,1]$ 连续, $f(0) = 0$、$f(1) = 1$ 且对于某个 $n \in \mathbb{N}$, 在 $[0,1]$ 上成立 $f^n(x) := (\underbrace{f \circ \cdots \circ f}_{n\text{次}})(x) \equiv x$, 则 $f(x) \equiv x$.

b) 若函数 $f: [0,1] \to [0,1]$ 连续且单调不减, 则对于任意的 $x \in [0,1]$, 或者 x 是不动点, 或者 $f^n(x)$ 趋于不动点, 这两种可能至少有一种实现 (这里 $f^n = f \circ \cdots \circ f$ 是 f 的 n 次迭代).

7. 设 $f: [0,1] \to \mathbb{R}$ 是连续函数, $f(0) = f(1)$. 证明

a) 对于任意的 $n \in \mathbb{N}$, 存在两端点在这个函数的图像上且长度等于 $\dfrac{1}{n}$ 的水平线段.

b) 若数 l 不是形如 $\dfrac{1}{n}$ 的数, 则存在所说的那样的函数, 在它的图像上已不能内接上长为 l 的水平线段

8. 以下述方式对于 $\delta > 0$ 定义的函数 $\omega(\delta)$, 叫做函数 $f: E \to \mathbb{R}$ 的连续模:

$$\omega(\delta) = \sup_{\substack{|x_1 - x_2| < \delta \\ x_1, x_2 \in E}} |f(x_1) - f(x_2)|.$$

这里, 上确界取遍集 E 的相互距离小于 δ 的一切可能的点对.

证明

a) 连续模是非减的非负的函数, 有极限[①]

$$\omega(+0) = \lim_{\delta \to 0+0} \omega(\delta).$$

b) 对于任意的 $\varepsilon > 0$, 存在 $\delta > 0$, 使得对于任意的点 x_1、$x_2 \in E$, 关系式 $|x_1 - x_2| < \delta$ 蕴含 $|f(x_1) - f(x_2)| < \omega(+0) + \varepsilon$.

c) 若 E 是闭区间、开区间或半开区间, 则对于函数 $f : E \to \mathbb{R}$ 的连续模成立着关系式

$$\omega(\delta_1 + \delta_2) \leqslant \omega(\delta_1) + \omega(\delta_2).$$

d) 在全直线上考察的函数 x 和 $\sin x^2$ 的连续模分别是在区域 $\delta > 0$ 中的函数 $\omega(\delta) = \delta$ 和常数函数 $\omega(\delta) = 2$.

e) 函数 f 在集 E 上一致连续当且仅当 $\omega(+0) = 0$.

9. 设 f 和 g 是定义在同一个集合 X 上的有界函数. 量

$$\Delta = \sup_{x \in X} |f(x) - g(x)|$$

称为有界函数f 和 g 之间的距离; 它表示在给定的集 X 上, 一个函数近似于另一个函数的好坏程度如何. 设 X 是闭区间 $[a, b]$. 证明, 如果 f、$g \in C[a, b]$, 那么, $\exists x_0 \in [a, b]$, 使 $\Delta = |f(x_0) - g(x_0)|$, 而对于任意的有界函数, 此事一般说来并不成立.

10. 设 $P_n(x)$ 是 n 阶多项式. 我们用多项式来逼近有界函数 $f : [a, b] \to \mathbb{R}$. 设

$$\Delta(P_n) = \sup_{x \in [a,b]} |f(x) - P_n(x)| \text{ 以及 } E_n(f) = \inf_{P_n} \Delta(P_n),$$

其中下确界取遍一切可能的 n 阶多项式. 如果 $\Delta(P_n) = E_n(f)$, 则多项式 P_n 叫做函数 f 的*最佳逼近多项式*.

证明

a) 存在零阶最佳逼近多项式$P_0(x) \equiv a_0$.

b) 当 P_n 是固定的多项式时, 在形如 $\lambda P_n(x)$ 的多项式 $Q_\lambda(x)$ 中间, 存在多项式 $Q_{\lambda_0}(x)$ 使得

$$\Delta(Q_{\lambda_0}) = \min_{\lambda \in \mathbb{R}} \Delta(Q_\lambda).$$

c) 如果 n 阶最佳逼近多项式存在, 则 $n+1$ 阶最佳逼近多项式也存在.

d) 对于任意的在闭区间上有界的函数和任意的 $n = 0, 1, 2, \cdots, n$ 阶最佳逼近多项式总存在.

11. 证明

a) 奇数阶实系数多项式至少有一个实根.

b) 若 $P_n(x)$ 是 n 阶多项式, 则函数 $\operatorname{sgn} P_n(x)$ 最多有 n 个间断点.

[①] 因此连续模通常对 $\delta \geqslant 0$ 来考察, 认为 $\omega(0) = \omega(+0)$.

c) 若在闭区间 $[a,b]$ 上有 $n+2$ 个点 $x_0 < x_1 < \cdots < x_{n+1}$ 使得量

$$\operatorname{sgn}[(f(x_i) - P(x_i))(-1)^i]$$

对于 $i = 0, 1, \cdots, n+1$ 为常数, 则 $E_n(f) \geqslant \min\limits_{0 \leqslant i \leqslant n+1} |f(x_i) - P_n(x_i)|$. 这是德拉瓦莱-布森① 定理. $E_n(f)$ 的定义见题 9.

12. a) 证明对于任意的 $n \in \mathbb{N}$, 闭区间 $[-1,1]$ 上定义的函数

$$T_n(x) = \cos(n \arccos x)$$

是 n 阶代数多项式(*切比雪夫多项式*).

b) 求出多项式 T_1、T_2、T_3、T_4 的明显的代数表达式并画出它们的图像.

c) 求出多项式 $T_n(x)$ 在闭区间 $[-1,1]$ 上的根以及量 $|T_n(x)|$ 在闭区间上达到最大值的点.

d) 证明在 x^n 项的系数为 1 的一切 n 阶多项式当中, 多项式 $T_n(x)$ 是唯一的与零偏差最小的多项式, 即

$$E_n(0) = \max_{|x| \leqslant 1} |T_n(x)|$$

($E_n(f)$ 的定义见题 9).

13. 设 $f \in C[a,b]$.

a) 证明, 若对于 n 阶多项式 $P_n(x)$ 存在 $n+2$ 个点 $x_0 < x_1 < \cdots < x_{n+1}$, 使得

$$f(x_i) - P_n(x_i) = (-1)^i \Delta(P_n) \cdot \alpha,$$

其中

$$\Delta(P_n) = \max_{x \in [a,b]} |f(x) - P_n(x)|,$$

而 α 是等于 1 或 -1 的常数 (这 $n+2$ 个点叫做*切比雪夫交错点*), 那么, $P_n(x)$ 同时也是函数 f 的唯一的 n 阶最佳逼近多项式 (见题 9).

b) 证明切比雪夫定理: n 阶多项式 $P_n(x)$ 是函数 $f \in C_{[a,b]}$ 的最佳逼近多项式当且仅当在闭区间 $[a,b]$ 上存在至少 $n+2$ 个切比雪夫交错点.

c) 证明对于间断函数上面的命题一般说来是不成立的.

d) 求出在闭区间 $[-1,2]$ 上函数 $|x|$ 的零阶和一阶最佳逼近多项式.

14. 在 §2 我们讨论过连续函数的局部性质. 本题把局部性质这个概念精确化.

我们认为两个函数 f 和 g 是在点 $a \in \mathbb{R}$ 处等价的, 如果存在该点的一个邻域 $U(a)$, 使 $\forall x \in U(a)$ 有 $f(x) = g(x)$. 两个函数间的这种关系显然是反身的, 对称的和传递的, 即确实是等价关系.

在点 a 处彼此等价的函数所成的族叫做在该点 a 处的*函数芽*. 如果只考察连续函数, 那么说的就是在点 a 处连续函数芽.

函数的局部性质就是指函数芽的性质.

①德拉瓦莱-布森(Ch.-J.de la Vallée Poussin) (1866—1962)—— 比利时数学家和力学家, 著有一本很好的分析教程.

a) 对同一点处的数值函数芽定义算术运算.

b) 证明, 在连续函数芽上的算术运算不越出这类芽的范围.

c) 顾及到 a 款和 b 款, 证明连续函数芽构成一个环——连续函数芽的环.

d) 称环 K 的子环 I 为环 K 的理想, 若环 K 的任意元与子环 I 的元的积都在 I 中, 试求点 a 处连续函数芽的环的理想.

15. 环的理想叫做极大的, 如果, 除自身以外, 它不含在任何更大的理想之中. 在闭区间上连续函数的集合 $C_{[a,b]}$ 关于数值函数的通常的加法和乘法运算构成一个环, 叫连续函数环, 求这个环的极大理想.

第五章　微分学

§1．可微函数

1．问题和引言　假设我们要根据牛顿[1]定律来解决二体开普勒[2]问题，就是说要弄清一个天体 m(行星) 相对于另一个天体 M(恒星) 的运动. 在运动平面内取以 M 为原点的笛卡儿坐标系 (图 13). 那么, m 在时刻 t 的位置可以用点 m 在这个坐标系中的数值坐标 $(x(t), y(t))$ 来表示. 我们要求函数 $x(t)$、$y(t)$.

m 相对于 M 的运动是受牛顿的两个著名的定律支配的:

一个是把力向量和由它引起的加速度向量通过正比例系数 m—— 物体的惯性质量[3]—— 联系起来的一般运动定律

$$\boldsymbol{ma} = \boldsymbol{F}; \tag{1}$$

另一个是万有引力定律, 它使我们能按公式

$$\boldsymbol{F} = G\frac{mM}{|\boldsymbol{r}|^3}\boldsymbol{r} \tag{2}$$

求出物体 m 和 M 相互间的引力, 其中 $\boldsymbol{r}$ 是以力所作用的物体为起点, 以另一个物体为终点的向量, $|\boldsymbol{r}|$ 是向量 $\boldsymbol{r}$ 的长度或者说是 m 和 M 之间的距离.

[1] 牛顿 (Newton) (1642—1727)—— 英国物理学家、力学家、天文学家和数学家, 最伟大的学者, 他阐述了经典力学的基本定律, 发现了万有引力定律,(与莱布尼茨同时) 奠定了微积分学的基础. 他同时代的人高度评价了他的贡献. 在他的坟碑上刻着:《这里安葬的只是牛顿的遗体》.

[2] 开普勒 (Kepler) (1571—1630)—— 著名的德国天文学家, 发现了行星运动定律 (开普勒定律).

[3] 我们用物体本身的记号表示质量, 这并不引起误会. 我们还注意到, 如果 $m \ll M$, 则可以认为坐标系是惯性系

知道了质量 m、M, 按照公式 (2) 不难通过物体 m 在时刻 t 的坐标 $x(t)$、$y(t)$ 表出方程式 (1) 的右端, 这已用到了所考察的运动的全部特征.

现在, 要想得到关于包含在方程 (1) 中的 $x(t)$、$y(t)$ 的关系式, 就必须学会通过函数 $x(t)$、$y(t)$ 来表示出方程式 (1) 的左端.

加速度是对于速度 $\boldsymbol{v}(t)$ 的变化的刻画, 简单而确切地说, 就是速度变化的速度; 因此, 为了解决我们的问题, 首先必须学会通过物体运动的矢径 $\boldsymbol{r}(t) = (x(t), y(t))$ 来计算物体在时刻 t 的速度 $\boldsymbol{v}(t)$.

于是, 我们就要定义并学会计算物体由运动定律 (1) 决定的瞬时速度.

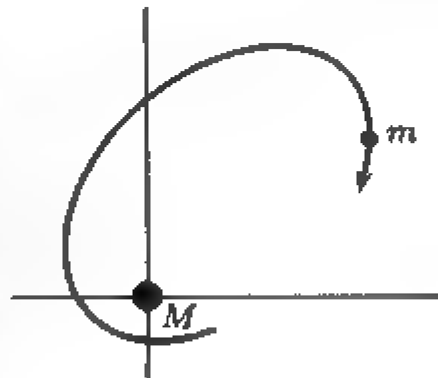


图 13

测量 —— 这就是与标准尺度进行比较. 在我们的情况下, 用来确定运动的瞬时速度的标准尺度是什么呢?

最简单的运动形式是自由物体完全按惯性运动的形式. 这是一种这样的运动, 在相同的时间间隔中, 物体在空间发生 (作为向量的) 相等的位移. 这就是所谓的匀速 (直线) 运动. 如果点作匀速运动, $\boldsymbol{r}(0)$ 和 $\boldsymbol{r}(1)$ 分别是它在时刻 $t = 0$ 和 $t = 1$ 相对于惯性坐标系的矢径, 那么在任意时刻都有

$$\boldsymbol{r}(t) - \boldsymbol{r}(0) = \boldsymbol{v} \cdot t, \tag{3}$$

其中 $\boldsymbol{v} = \boldsymbol{r}(1) - \boldsymbol{r}(0)$. 于是, 位移 $\boldsymbol{r}(t) - \boldsymbol{r}(0)$ 在最简单的情况下是时间的*线性函数*, 在这种情况下, 在单位时间内的位移向量 $\boldsymbol{v}$ 起着位移 $\boldsymbol{r}(t) - \boldsymbol{r}(0)$ 和时间 t 之间的比例因子的作用. 这个向量就叫做匀速运动的速度. 至于运动是直线的这一结论, 可以从它的轨迹的参数方程:

$$\boldsymbol{r}(t) = \boldsymbol{r}(0) + \boldsymbol{v} \cdot t$$

是直线方程看出来 (见解析几何教程).

于是我们知道了由公式 (3) 给出的匀速直线运动的速度 $\boldsymbol{v}$. 根据惯性定律, 如果物体不受外力, 它就做匀速直线运动. 这就意味着, 如果在时刻 t 取消 M 对物体 m 的作用, 那么后者将以某个确定的速度匀速地继续自己的运动. 于是, 自然认为正是这个速度是我们的物体在时刻 t 的 (瞬时) 速度.

不过, 如果没有下面我们立即就来讨论的最重要的情况的话, 瞬时速度的这种定义就会只是纯粹的抽象, 给不出计算这个量的任何办法.

当滞留在我们已进入的这个 (从逻辑上说好像是有"漏洞"的) 范围内时, 写下运动方程 (1), 跟着就解释瞬时速度和加速度是什么, 那么, 我们毕竟会发现, 在这些概念的最一般的表达形式下, 从方程式 (1) 还是可以作出如下有启发性的论断. 如果力不存在, 即 $\boldsymbol{F} = \boldsymbol{0}$, 那么, 加速度 也等于零. 但若速度 $\boldsymbol{v}(t)$ 的变化速度 $\boldsymbol{a}(t)$ 等于零, 则显然速度 $\boldsymbol{v}(t)$ 大概不随时间变化. 于是我们得到惯性定律, 按此定律, 自由物体实际上在空间中以不随时间变化的速度运动.

从方程 (1) 还看出, 数量上有界的力只可能产生数量上有界的加速度. 但若在时间间隔 $[0, t]$ 内, 某个量 $P(t)$ 的变化速度的绝对值不超过某个常数 c, 则按我们的表达方式, 量 P 在时间 t 内的改变 $|P(t) - P(0)|$ 不超过 $c \cdot t$, 也就是说, 在这种情况下, 在小的时间间隔中量只做小的改变 (在任何情况下函数 $P(t)$ 都是连续的). 这就意味着, 实际的力学系统在小的时间间隔中它的参量只做小的改变.

特别地, 物体 m 的速度 $\boldsymbol{v}(t)$ 在接近时刻 t_0 的一切时刻 t 都应该是接近于我们企图确定的值 $\boldsymbol{v}(t_0)$ 的. 但在这样的情况下, 运动本身在时刻 t_0 的小的邻域内应该与以速度 $\boldsymbol{v}(t_0)$ 所作的匀速运动差别不大, 并且我们离开 t_0 越近, 这个差别就应越小.

如果我们通过天文望远镜拍摄下天体 m 的轨迹的话, 那么, 依赖于望远镜的性能, 我们所得大致如下:

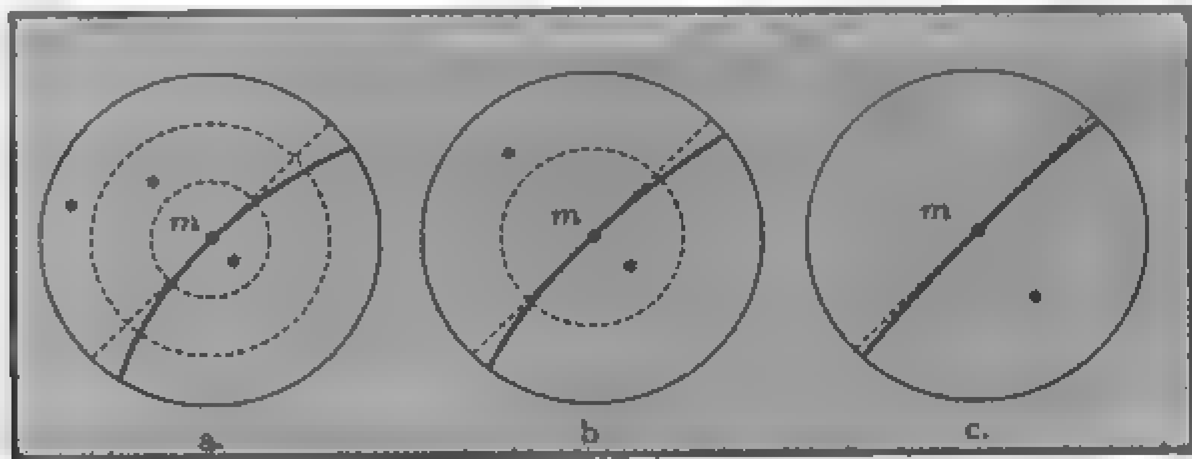


图 14

照片 c 上显示的一段轨迹相应于一段相当小的时间间隔, 以致于在这段时间中已难于把真正的轨迹同直线区别开来, 因为这段轨迹确实像是直线, 而运动就像是匀速直线似的. 顺便说一下, 由这一考察, 可以断定, 解决了瞬时速度的定义问题 (速度是向量), 我们就同时解决了曲线 (在上述情况下的曲线就是运动轨迹) 的切线的定义和求法这样一个纯几何的问题.

于是我们发现, 在我们的问题中, 当 t 接近 t_0 时, 应有 $\boldsymbol{v}(t) \approx \boldsymbol{v}(t_0)$, 即当 $t \to t_0$ 时, $\boldsymbol{v}(t) \to \boldsymbol{v}(t_0)$, 也就是当 $t \to t_0$ 时, $\boldsymbol{v}(t) = \boldsymbol{v}(t_0) + o(1)$. 此时还应有

$$\boldsymbol{r}(t) - \boldsymbol{r}(t_0) \approx \boldsymbol{v}(t_0)(t - t_0)$$

当 t 接近 t_0 时成立, 更精确地说, 当 $t \to t_0$ 时, 位置的改变量 $\boldsymbol{r}(t) - \boldsymbol{r}(t_0)$ 等价于 $\boldsymbol{v}(t_0)(t - t_0)$, 或

$$\boldsymbol{r}(t) - \boldsymbol{r}(t_0) = \boldsymbol{v}(t_0)(t - t_0) + o(\boldsymbol{v}(t_0)(t - t_0)), \tag{4}$$

其中 $o(\boldsymbol{v}(t_0)(t - t_0))$ 是修正向量, 它的值当 $t \to t_0$ 时比向量 $\boldsymbol{v}(t_0)(t - t_0)$ 的值趋于零更快. 这里, 自然应该对于 $\boldsymbol{v}(t_0) = \mathbf{0}$ 的情形作出说明. 为了不把这种情形从一般

的考察中排除出去, 宜于指出

$$|\boldsymbol{v}(t_0)(t-t_0)| = |\boldsymbol{v}(t_0)||t-t_0|^{①}.$$

这样一来, 如果 $|\boldsymbol{v}(t_0)| \neq 0$, 那么, 量 $|\boldsymbol{v}(t_0)(t-t_0)|$ 与 $|t-t_0|$ 同阶, 从而 $o(\boldsymbol{v}(t_0)(t-t_0)) = o(t-t_0)$. 这就是说, 代替 (4) 可以写关系式

$$\boldsymbol{r}(t) - \boldsymbol{r}(t_0) = \boldsymbol{v}(t_0)(t-t_0) + o(t-t_0), \tag{5}$$

这个式子是不排除 $\boldsymbol{v}(t_0) = \boldsymbol{0}$ 的情形的.

这样一来, 从关于速度的最一般的然而也许是含糊的表述我们过渡到了速度应该满足的关系式 (5). 但是量 $\boldsymbol{v}(t_0)$ 从 (5) 式单值求出:

$$\boldsymbol{v}(t_0) = \lim_{t\to t_0} \frac{\boldsymbol{r}(t) - \boldsymbol{r}(t_0)}{t-t_0}, \tag{6}$$

因此, 无论是基本关系式 (5) 本身还是与它等价的关系式 (6), 现在都可取作物体 m 在时刻 t_0 的瞬时速度 $\boldsymbol{v}(t_0)$ 的定义.

我们不打算现在就对向量值函数的极限问题做详细讨论, 而只限于把这个问题归结为已经详细考察过的实值函数的极限的情形. 因为向量 $\boldsymbol{r}(t) - \boldsymbol{r}(t_0)$ 有坐标 $(x(t)-x(t_0), y(t)-y(t_0))$, 所以

$$\frac{\boldsymbol{r}(t) - \boldsymbol{r}(t_0)}{t-t_0} = \left(\frac{x(t)-x(t_0)}{t-t_0}, \frac{y(t)-y(t_0)}{t-t_0}\right).$$

这就意味着, 如果认为向量间的接近是指的它们的坐标相接近的话, 则 (6) 中的极限应该理解为:

$$\boldsymbol{v}(t_0) = \lim_{t\to t_0} \frac{\boldsymbol{r}(t) - \boldsymbol{r}(t_0)}{t-t_0} = \left(\lim_{t\to t_0} \frac{x(t)-x(t_0)}{t-t_0}, \lim_{t\to t_0} \frac{y(t)-y(t_0)}{t-t_0}\right),$$

而 (5) 中的 $o(t-t_0)$ 应该理解作一个这样的依赖于 t 的向量, 它使得当 $t \to t_0$ 时向量 $\dfrac{o(t-t_0)}{t-t_0}$ 趋于零 (即按坐标趋于零).

最后, 我们发现, 若 $\boldsymbol{v}(t_0) \neq \boldsymbol{0}$, 则方程

$$\boldsymbol{r}(t) - \boldsymbol{r}(t_0) = \boldsymbol{v}(t_0)(t-t_0) \tag{7}$$

给出一条直线, 根据上述情况, 这条直线应该当作是轨迹在点 $(x(t_0), y(t_0))$ 处的切线.

于是, 由线性关系式 (7) 给出的匀速直线运动的速度就是确定运动速度的标准尺度. 标准的运动方程 (7) 适合于所研究的运动, 因为它是关系式 (5) 的要求的. 使 (5) 成立的值 $\boldsymbol{v}(t_0)$ 可以通过极限过程 (6) 求出, 叫做时刻 t_0 时运动的速度, 在经典

---

①这里 $|t-t_0|$ 是数 $t-t_0$ 的模, 而 $|\boldsymbol{v}|$ 是向量 $\boldsymbol{v}$ 的模即长度.

力学中研究的用定律 (1) 描述的运动, 应该允许与这样的标准尺度来比较, 即应该允许做 (5) 中所述的线性近似.

如果 $\boldsymbol{r}(t)=(x(t),y(t))$ 是动点 m 在时刻 t 的矢径, $\dot{\boldsymbol{r}}(t)=(\dot{x}(t),\dot{y}(t))=\boldsymbol{v}(t)$ 是 $\boldsymbol{r}(t)$ 在时刻 t 的变化速度向量, 而 $\ddot{\boldsymbol{r}}(t)=(\ddot{x}(t),\ddot{y}(t))=\boldsymbol{a}(t)$ 是 $\boldsymbol{v}(t)$ 在时刻 t 的变化速度向量即加速度, 那么方程式 (1) 可以写成

$$m\ddot{\boldsymbol{r}}(t)=\boldsymbol{F}(t).$$

由此, 对于在引力场中的运动, 我们得到方程 (1) 的坐标形式

$$\begin{cases}\ddot{x}(t)=-GM\dfrac{x(t)}{[x^2(t)+y^2(t)]^{3/2}},\\ \ddot{y}(t)=-GM\dfrac{y(t)}{[x^2(t)+y^2(t)]^{3/2}}.\end{cases}\tag{8}$$

这就是我们原来的问题的精确的数学写法. 因为我们知道怎样由 $\boldsymbol{r}(t)$ 求 $\dot{\boldsymbol{r}}(t)$ 以及 $\ddot{\boldsymbol{r}}(t)$, 所以我们现在已经面临着这样一个问题, 有什么函数对 $(x(t),y(t))$ 能给出物体 m 绕 M 的运动吗? 为此, 必须求出 $\ddot{x}(t)$、$\ddot{y}(t)$ 并检验关系式 (8) 是否成立. 方程组 (8) 是所谓微分方程组的一个例子. 暂时我们只会验算某一组函数是不是方程组的解. 怎样求解, 或者更正确些说, 怎样研究微分方程的解的性质, 这些问题在分析学的一个专门的分支 —— 微分方程论中研究, 此刻我们已能明白, 这是一个非常重要的学科.

求向量变化速度的运算, 如已表明的, 归结为求某些数值函数 —— 向量的坐标的变化速度. 因此, 必须首先学会在最简单的实自变量的实值函数的情况下自如地完成这种运算. 我们现在就来做这件事.

2. 在一点处可微的函数 我们从两个预备性的定义开始, 稍后再将它们做某种程度的精确化.

定义 0_1 定义在集 $E\subset\mathbb{R}$ 上的函数 $f:E\to\mathbb{R}$ 叫做在 E 的极限点 $a\in E$ 处可微, 如果存在一个关于自变量的增量 $x-a$ 的线性函数 $A\cdot(x-a)$, 使得函数的增量 $f(x)-f(a)$ 表成

$$f(x)-f(a)=A(x-a)+o(x-a),\text{当}x\to a,x\in E.\tag{9}$$

换句话说, 函数在点 a 处可微是指它的值在所研究的点的邻域内的变化近似是线性的, 这种近似的误差同点 a 的变化量 $x-a$ 相比是无穷小量.

定义 0_2 (9) 中的线性函数 $A(x-a)$ 叫做函数 f 在点 a 的微分.

函数在一点的微分是单值确定的, 因为从 (9) 推出

$$\lim_{E\ni x\to a}\frac{f(x)-f(a)}{x-a}=\lim_{E\ni x\to a}\left(A+\frac{o(x-a)}{x-a}\right)=A,$$

而根据极限的唯一性, 数 A 是单值确定的.

定义 1 量

$$f'(a)=\lim_{E\ni x\to a}\frac{f(x)-f(a)}{x-a} \tag{10}$$

叫做函数 f 在点 a 处的导数.

关系式 (10) 可以改写成等价形式

$$\frac{f(x)-f(a)}{x-a}=f'(a)+\alpha(x),$$

其中当 $x\to a, x\in E$ 时, $\alpha(x)\to 0$, 这同样等价于关系式

$$f(x)-f(a)=f'(a)(x-a)+o(x-a)\text{当}x\to a, x\in E. \tag{11}$$

因此, 函数可微等价于在相应的点处它有导数存在.

如果把这些定义同第一段中所述内容相比照, 则可断定, 导数刻画函数在所考察的点处的变化速度, 而微分则提供了函数在所考察的点的邻域内的最佳线性近似.

如果函数 $f: E\to\mathbb{R}$ 在集 E 的不同的点处可微, 那么当从一点移到另一点时, 在 (9) 中无论是量 A 还是函数 $o(x-a)$ 都可能改变 (在 (11) 中我们已经清楚地觉察到了这一事实). 这种状况确实应该在可微函数的定义中指出, 我们现在就给出这一基本定义的完整写法.

定义 2 定义在集 $E\subset\mathbb{R}$ 上的函数 $f: E\to\mathbb{R}$ 叫做在集 E 的极限点 $x\in E$ 处是可微的, 如果

$$\boxed{f(x+h)-f(x)=A(x)h+\alpha(x;h),} \tag{12}$$

其中 $h\to A(x)h$ 是关于 h 的线性函数, 而 $\alpha(x;h)$ 当 $h\to 0, x+h\in E$ 时, 等于 $o(h)$.

量

$$\Delta x(h):=(x+h)-x=h,$$

和

$$\Delta f(x;h):=f(x+h)-f(x),$$

分别叫做*自变量的增量*和*函数* (相应于自变量的这个增量) *的增量*.

它们常常用 (其实并不完全合理) 自己作为 h 的函数的记号 Δx 和 $\Delta f(x)$ 表示.

这样一来, 说函数在一点处可微, 指的是它在这点处的增量作为自变量增量 h 的函数近似于一个线性函数, 其误差与自变量增量的比当 $h\to 0$ 时是无穷小量.

定义 3 定义 2 中关于 h 的线性函数 $h\to A(x)h$ 叫做函数 $f: E\to\mathbb{R}$ 在点 $x\in E$ 处的*微分*, 并用符号 $df(x)$ 或 $Df(x)$ 来表示.

于是, $df(x)(h) = A(x)h$.

从定义 2、3, 我们有

$$\Delta f(x;h) - df(x)(h) = \alpha(x;h),$$

并且当 $h \to 0, x+h \in E$ 时, $\alpha(x;h) = o(h)$, 就是说, 由自变量的增量 h 引起的函数增量与线性函数 $df(x)$ 在同一个 h 处的值之间的差是关于 h 的高于一阶的无穷小量.

由于这个缘故, 我们说,**微分是函数增量的线性 (主要) 部分**.

从关系式 (12) 和定义 1 可以推出,

$$A(x) = f'(x) = \lim_{\substack{h \to 0 \\ x+h, x \in E}} \frac{f(x+h) - f(x)}{h},$$

因此微分可以写成

$$df(x)(h) = f'(x)h. \tag{13}$$

特别地, 如果 $f(x) \equiv x$, 则显然 $f'(x) \equiv 1$ 且

$$dx(h) = 1 \cdot h = h,$$

因此有时也说,"自变量的微分就是它的增量".

注意到这个等式, 从 (13) 得到

$$df(x)(h) = f'(x)dx(h), \tag{14}$$

即

$$df(x) = f'(x)dx. \tag{15}$$

等式 (15) 必须理解作 h 的函数的等式.

从 (14) 得到

$$\frac{df(x)(h)}{dx(h)} = f'(x), \tag{16}$$

即函数 $\frac{df(x)}{dx}$ (函数 $df(x)$ 和 dx 的比) 是常数并且等于 $f'(x)$. 由于这个缘故, 人们常常根据莱布尼茨的办法, 用记号 $\frac{df(x)}{dx}$ 来标记导数. 这种记法与后来拉格朗日[①]提出的记号 $f'(x)$ 均为人们所使用.

在力学中, 除了上述记号外, 还用记号 $\dot{\varphi}(t)$ 来表示函数 $\varphi(t)$ 关于时间 t 的导数 (读作 "φ 点 t").

---

①拉格朗日 (Lagrange) (1736—1813)—— 著名的法国数学家和力学家.

3. 切线; 导数和微分的几何意义 设 $f: E \to \mathbb{R}$ 是定义在集 $E \subset \mathbb{R}$ 上的函数, 而 x_0 是集 E 的一个固定的极限点. 我们希望挑选一个常数 c_0, 使得它比别的任何常数都能更好地刻画函数在点 x_0 的邻域内的性态. 精确地说, 我们希望差 $f(x) - c_0$ 当 $x \to x_0, x \in E$ 时, 与任何非零常数相比都是无穷小量, 即

$$f(x) = c_0 + o(1) \quad 当 \quad x \to x_0, x \in E. \tag{17}$$

这个关系式等价于

$$\lim_{E \ni x \to x_0} f(x) = c_0.$$

特别地, 如果函数在点 x_0 处连续, 则

$$\lim_{E \ni x \to x_0} f(x) = f(x_0),$$

于是当然有 $c_0 = f(x_0)$.

现在我们要求选择一个函数 $c_0 + c_1(x - x_0)$, 使

$$f(x) = c_0 + c_1(x - x_0) + o(x - x_0) \quad 当 \quad x \to x_0, x \in E. \tag{18}$$

显然, 这是前述问题的推广, 因为可以把公式 (18) 改写成

$$f(x) = c_0 + o((x - x_0)^0) \quad 当 \quad x \to x_0, x \in E.$$

让 $x \to x_0, x \in E$, 从 (18) 式直接推出 $c_0 = \lim\limits_{E \ni x \to x_0} f(x)$, 若函数在这点连续, 则 $c_0 = f(x_0)$.

如果 c_0 找到了, 那么从 (18) 推出

$$c_1 = \lim_{E \ni x \to x_0} \frac{f(x) - c_0}{x - x_0}.$$

而一般说来, 假如我们找到了一个多项式

$$P_n(x_0; x) = c_0 + c_1(x - x_0) + \cdots + c_n(x - x_0)^n,$$

使得

$$f(x) = c_0 + c_1(x - x_0) + \cdots + c_n(x - x_0)^n + o((x - x_0)^n) 当 x \to x_0, x \in E. \tag{19}$$

那么, 我们就可以顺次完全单值地求出

$$c_0 = \lim_{E \ni x \to x_0} f(x), \quad c_1 = \lim_{E \ni x \to x_0} \frac{f(x) - c_0}{x - x_0},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$c_n = \lim_{E \ni x \to x_0} \frac{f(x) - [c_0 + \cdots + c_{n-1}(x - x_0)^{n-1}]}{(x - x_0)^n},$$

如果上述所有极限都存在的话; 而在相反的情况下, 条件 (19) 不成立, 从而问题无解.

如果函数 f 在点 x_0 连续, 那么, 如已指出的, 从 (18) 推出 $c_0 = f(x_0)$, 从而我们得到关系式

$$f(x) - f(x_0) = c_1(x - x_0) + o(x - x_0) \text{ 当 } x \to x_0, x \in E,$$

这与函数 $f(x)$ 在点 x_0 可微的条件等价.

由此求得

$$c_1 = \lim_{E \ni x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0).$$

这样一来就证明了

命题 1 *在集 $E \subset \mathbb{R}$ 的极限点 x_0 处连续的函数 $f : E \to \mathbb{R}$ 使线性近似式 (18) 成立的充分必要条件是它在这个点处可微.*

函数

$$\varphi(x) = c_0 + c_1(x - x_0) \tag{20}$$

当 $c_0 = f(x_0)$ 且 $c_1 = f'(x_0)$ 时, 是满足关系式 (18) 的唯一的形如 (20) 的函数.

于是函数

$$\varphi(x) = f(x_0) + f'(x_0)(x - x_0) \tag{21}$$

依下述意义给出了函数 f 在点 x_0 的邻域中的最佳线性近似: 对于任何其他形如 (20) 的函数 $\varphi(x)$, 当 $x \to x_0, x \in E$ 时, $f(x) - \varphi(x) \neq o(x - x_0)$.

函数 (21) 的图像是过点 $(x_0, f(x_0))$、斜率为 $f'(x_0)$ 的直线

$$y - f(x_0) = f'(x_0)(x - x_0). \tag{22}$$

由于直线 (22) 提供了函数 $y = f(x)$ 的图像在点 $(x_0, f(x_0))$ 的邻域内的最佳可能的线性近似, 那么, 自然会接受.

定义 4 如果函数 $f : E \to \mathbb{R}$ 定义在集 $E \subset \mathbb{R}$ 且在点 $x_0 \in E$ 处可微, 那么, 方程 (22) 给出的直线就叫做这个函数的图像在点 $(x_0, f(x_0))$ 处的切线.

图 15 解释了至此我们已经引入的与函数在一点处的可微性相关的所有基本概念: 自变量的增量和与它相应的函数增量及微分的值; 在图上画出了函数的图像、图像在点 $P_0 = (x_0, f(x_0))$ 处的切线以及用以进行比较的任意的一条过 P_0 和函数图像上某点 $P \neq P_0$ 的直线 (通常叫做割线).

定义 4 的拓广是

定义 5 如果映射 $f : E \to \mathbb{R}$, $g : E \to \mathbb{R}$ 都在集 $E \subset \mathbb{R}$ 的极限点 $x_0 \in E$ 处连续且当 $x \to x_0$、$x \in E$ 时 $f(x) - g(x) = o((x - x_0)^n)$, 那么就说 f 和 g 在点 x_0 处 n阶相切 (或者更确切地说不低于 n 阶相切).

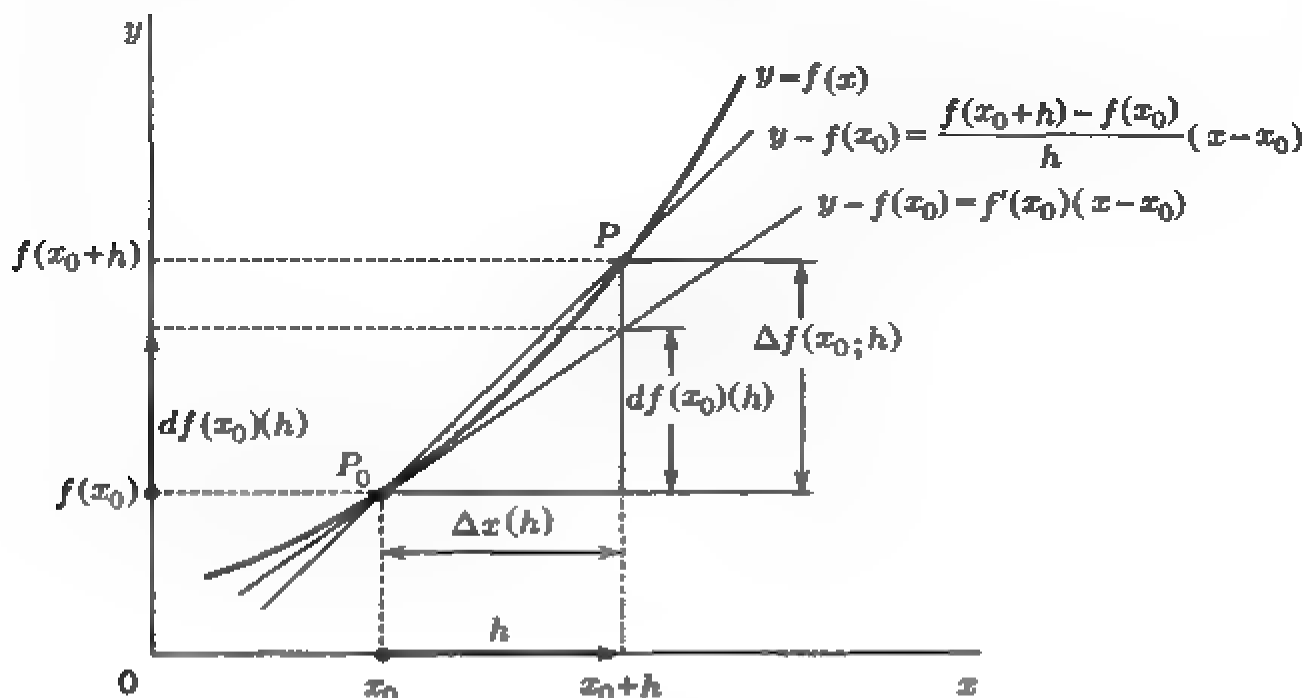


图 15

当 $n=1$ 时, 说 f 和 g 在点 x_0 处彼此相切.

按照定义 5, 映射 (21) 在点 x_0 处与在该点可微的映射 $f: E \to \mathbb{R}$ 相切.

现在也可以说, 关系式 (19) 中的多项式 $P(x_0;x)=c_0+c_1(x-x_0)+\cdots+c_n(x-x_0)^n$ 与函数 f 不低于 n 阶相切.

可以把数 $h=x-x_0$, 即自变量的增量, 看作是一个从 x_0 到 $x=x_0+h$ 的位移的向量. 把这样的向量的全体记作[①] $T\mathbb{R}(x_0)$ 或者 $T\mathbb{R}_{x_0}$. 类似地, 用 $T\mathbb{R}(y_0)$ 或 $T\mathbb{R}_{y_0}$ 来表示从点 y_0 沿 y 轴的位移向量的全体 (见图 15). 那么, 从微分的定义看出, 由微分 $h \mapsto f'(x_0)h = df(x_0)(h)$ 给出的映射

$$df(x_0): T\mathbb{R}(x_0) \to T\mathbb{R}(f(x_0)) \tag{23}$$

与可微函数的增量所给出的映射

$$h \mapsto f(x_0+h)-f(x_0)=\Delta f(x_0;h) \tag{24}$$

是相切的.

我们指出 (见图 15), 如果说映射 (24) 是函数 $y=f(x)$ 的图像的纵坐标当自变量从点 x_0 变到点 x_0+h 时的增量的话, 则微分 (23) 给出了函数图像的切线的纵坐标在自变量的同一增量 h 之下的增量.

4. 坐标系的作用 人们对切线的解析的定义 4 可能产生某种莫明的不满. 我们将尽力说清是什么东西令人不满. 不过我们先从更加侧重几何的角度指出曲线在它的一点 P_0 处的切线的结构 (见图 15).

① 这与通用的记号 $T_{x_0}\mathbb{R}$ 或 $T_{x_0}(\mathbb{R})$ 稍有不同.

取曲线的异于 P_0 的任意的点 P. 由两点 P_0、P 决定的直线, 已经指出过, 叫做曲线的割线. 现在让点 P 沿着曲线趋于点 P_0, 如果这时割线趋于某个极限位置, 那么割线的这个极限位置就是曲线在点 P_0 处的切线.

切线的这种全凭直观的定义, 此刻对于我们来说是不能接受的, 因为我们不知道曲线是什么, 不知道什么叫做"点沿着曲线趋于另一点", 最后, 我们也不知道应该怎样理解"割线的极限位置".

我们此刻不想把这些概念精确化, 而是集中注意关于切线的这两种定义之间的基本差别. 第二个定义是纯几何的, 它与坐标系的选取没有关系. 而在第一种情形时, 我们考察的曲线都是可微函数在某个坐标系中图像, 定义的切线都是这种形式的曲线的切线. 自然可能发生这样的问题, 当着在另一个坐标系中表出这条曲线时, 相应的函数会不会变成不可微的, 或者虽然可微, 但新的计算的结果所得到的切线却是另一条直线.

这个不变性问题, 即与坐标系的无关性的问题, 当一个概念是借助于某个坐标系引入的时候, 总是要发生的.

对于速度概念也同样有这个问题. 速度的概念我们已在第 1 段中讨论过, 并且顺便说一下, 正像已经见到的, 这个概念包含了切线的概念.

点、向量、直线等等在不同的坐标系中有不同的数量表示 (点的坐标、向量的坐标、直线的方程). 不过, 只要知道了联系两个坐标系的公式, 就一定可以弄清楚两个同一类型的数字表示是不是同一个几何对象在不同的坐标系中的记法. 直觉向我们暗示, 第 1 段中所描述的定义速度的手续将产生同一个向量, 而与在其中进行计算的坐标系无关, 到研究多元函数的时候, 我们将详细地讨论类似的问题. 而在下一节我们就来验证速度的定义关于不同的坐标系的不变性.

在着手考察具体的例子之前我们先总结一下.

我们已经接触到对运动物体的瞬时速度进行数学描述这样一个问题.

这个问题归结到对于给定的函数在所研究的点的邻域内用线性函数来近似的问题, 从几何学上来说, 归结到切线的概念. 描述实际力学系统的运动的函数, 常假定是允许做这样的线性近似的.

这样就自然地从所有函数中分出了一类可微函数.

函数在一点处的微分的概念是作为线性映射引人的, 它的定义域是从所考察的点出发的位移的集合, 它描述可微函数的增量在所考察的点的邻域内的性态, 其误差与位移的大小相比是无穷小量.

微分 $df(x_0)(h) = f'(x_0)h$ 完全由函数 f 在点 x_0 的导数$f'(x_0)$ 决定, 此数可以经极限过渡求得

$$f'(x_0) = \lim_{E \ni x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

导数的物理意义是量 $f(x)$ 在时刻 x_0 的变化速度; 导数的几何意义是函数 $y = f(x)$ 的图像在点 $(x_0, f(x_0))$ 处的切线的斜率.

5. 一些例子

例 1 设 $f(x)=\sin x$. 证明 $f'(x)=\cos x$.

$$◀ \lim_{h\to 0}\frac{\sin(x+h)-\sin x}{h}=\lim_{h\to 0}\frac{2\sin\frac{h}{2}\cdot\cos\left(x+\frac{h}{2}\right)}{h}$$

$$=\lim_{h\to 0}\cos\left(x+\frac{h}{2}\right)\cdot\lim_{h\to 0}\frac{\sin\frac{h}{2}}{\frac{h}{2}}=\cos x. ▶$$

我们使用了关于乘积的极限的定理, 函数 $\cos x$ 的连续性, 当 $t\to 0$ 时的等价关系 $\sin t\sim t$ 以及关于复合函数的极限的定理.

例 2 证明 $\cos' x=-\sin x$.

$$◀ \lim_{h\to 0}\frac{\cos(x+h)-\cos x}{h}=\lim_{h\to 0}\frac{-2\sin\frac{h}{2}\cdot\sin\left(x+\frac{h}{2}\right)}{h}$$

$$=-\lim_{h\to 0}\sin\left(x+\frac{h}{2}\right)\lim_{h\to 0}\frac{\sin\frac{h}{2}}{\frac{h}{2}}=-\sin x. ▶$$

例 3 证明, 若 $f(t)=r\cos\omega t$ 则 $f'(t)=-r\omega\sin\omega t$.

$$◀ \lim_{h\to 0}\frac{r\cos\omega(t+h)-r\cos\omega t}{h}=r\lim_{h\to 0}\frac{-2\sin\left(\omega\frac{h}{2}\right)\sin\omega\left(t+\frac{h}{2}\right)}{h}$$

$$=-r\lim_{h\to 0}\sin\omega\left(t+\frac{h}{2}\right)\lim_{h\to 0}\frac{\sin\left(\omega\frac{h}{2}\right)}{\left(\omega\frac{h}{2}\right)}\cdot\omega$$

$$=-r\omega\sin\omega t. ▶$$

例 4 若 $f(t)=r\sin\omega t$ 则 $f'(t)=r\omega\cos\omega t$.

◀ 证明类似例 1 和例 3. ▶

例 5 *质点的瞬时速度和瞬时加速度.* 设质点在平面上运动, 并且, 在固定的坐标系中它的运动规律是用时间的可微函数

$$x=x(t),\quad y=y(t)$$

来描述的, 亦即用向量 $\boldsymbol{r}(t)=(x(t),y(t))$ 来描述. 在本节第 1 段中我们已经弄清楚了点在时刻 t 的速度是向量

$$\boldsymbol{v}(t)=\dot{\boldsymbol{r}}(t)=(\dot{x}(t),\dot{y}(t)),$$

其中 $\dot{x}(t)$、$\dot{y}(t)$ 是函数 $x(t)$、$y(t)$ 关于时间 t 的导数.

加速度 $\boldsymbol{a}(t)$ 是向量 $\boldsymbol{v}(t)$ 的变化速度, 因此

$$\boldsymbol{a}(t)=\dot{\boldsymbol{v}}(t)=\ddot{\boldsymbol{r}}(t)=(\ddot{x}(t),\ddot{y}(t)),$$

其中 $\ddot{x}(t), y(t)$ 是函数 $\dot{x}(t)$、$\dot{y}(t)$ 关于 t 的导数, 也叫函数 $x(t)$、$y(t)$ 的第二阶导数.

因此, 根据问题的物理意义, 描述质点运动的函数 $x(t)$、$y(t)$ 应该既具有一阶, 也具有二阶导数.

特别地, 我们来考察点沿着半径为 r 的圆周的匀速运动. 设 ω 是点的角速度, 即点在单位时间内移过的圆周角的值.

在笛卡儿坐标系中 (根据函数 $\cos x$、$\sin x$ 的定义) 这个运动表示成

$$\boldsymbol{r}(t)=(r\cos(\omega t+\alpha), r\sin(\omega t+\alpha)),$$

而若 $\boldsymbol{r}(0)=(r,0)$ 则写成

$$\boldsymbol{r}(t)=(r\cos\omega r, r\sin\omega t).$$

为了简化记法, 将认为 $\boldsymbol{r}(0)=(r,0)$, 这并不损害后面结果的一般性.

那么, 根据例 3 和例 4 的结果, 有

$$\boldsymbol{v}(t)=\dot{\boldsymbol{r}}(t)=(-r\omega\sin\omega t, r\omega\cos\omega t).$$

不出所料, 从内积的计算

$$(\boldsymbol{v}(t),\boldsymbol{r}(t))=-r^2\omega\sin\omega t\cos\omega t+r^2\omega\cos\omega t\sin\omega t=0,$$

得到, 速度向量 $\boldsymbol{v}(t)$ 与矢径 $\boldsymbol{r}(t)$ 正交且与圆周的切线同一指向.

其次, 对于加速度有

$$\boldsymbol{a}(t)=\dot{\boldsymbol{v}}(t)=\ddot{\boldsymbol{r}}(t)=(-r\omega^2\cos\omega t,-r\omega^2\sin\omega t),$$

即 $\boldsymbol{a}(t)=-\omega^2\boldsymbol{r}(t)$, 因此, 加速度确实是指向中心的, 因为它与向量 $\boldsymbol{r}(t)$ 反方向.

还有

$$|\boldsymbol{a}(t)|=\omega^2|\boldsymbol{r}(t)|=\omega^2 r=\frac{|\boldsymbol{v}(t)|^2}{r}=\frac{v^2}{r},$$

其中 $v=|\boldsymbol{v}(t)|$.

从这个公式出发, 我们来计算, 例如, 地球的低卫星的速度值. 在这种情况下 r 等于地球的半径, 即 $r\approx 6400$ 公里, 而 $|\boldsymbol{a}(t)|=g$, 其中 $g\approx 10$ 米/秒 ^{2} 是地球表面自由落体的加速度.

于是 $v^2=|\boldsymbol{a}(t)|r\approx 10$ 米/秒 $^2\times 64\times 10^5$米$=64\times 10^6$(米/秒)^2. 从而 $v\approx 8\times 10^3$ 米/秒.

例 6 抛物镜的光学性质. 考察 (图 16) 抛物线 $y=\frac{1}{2p}x^2(p>0)$, 并做它在点 $(x_0,y_0)=\left(x_0,\frac{1}{2p}x_0^2\right)$ 处的切线.

由于 $f(x)=\frac{1}{2p}x^2$, 所以

$$f'(x_0)=\lim_{x\to x_0}\frac{\frac{1}{2p}x^2-\frac{1}{2p}x_0^2}{x-x_0}=\frac{1}{2p}\lim_{x\to x_0}(x+x_0)=\frac{1}{p}x_0.$$

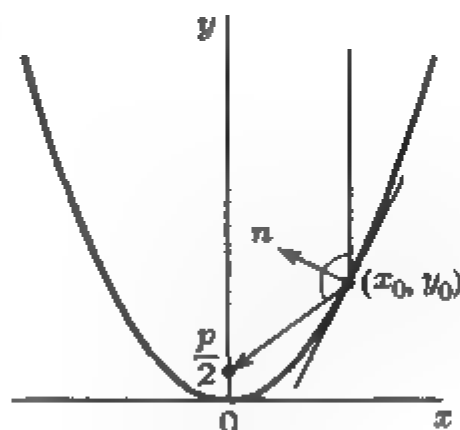


图 16

这就是说, 所求的切线的方程是

$$y-\frac{1}{2p}x_0^2=\frac{1}{p}x_0(x-x_0),$$

或

$$\frac{1}{p}x_0(x-x_0)-(y-y_0)=0, \tag{25}$$

其中 $y_0=\frac{1}{2p}x_0^2$.

从最后这个方程见到, 向量 $\boldsymbol{n}=\left(-\frac{1}{p}x_0,1\right)$ 与直线 (25) 正交. 我们来证明向量 $\boldsymbol{e}_y=(0,1)$ 和 $\boldsymbol{e}_f=\left(-x_0,\frac{p}{2}-y_0\right)$ 与 $\boldsymbol{n}$ 成相等的夹角. 向量 $\boldsymbol{e}_y$ 是轴 Oy 方向的单位向量, 而 $\boldsymbol{e}_f$ 是从切点 $(x_0,y_0)=\left(x_0,\frac{1}{2p}x_0^2\right)$ 指向抛物线的焦点 $\left(0,\frac{p}{2}\right)$ 的向量. 于是

$$\cos\widehat{\boldsymbol{e}_y\boldsymbol{n}}=\frac{\langle\boldsymbol{e}_y,\boldsymbol{n}\rangle}{|\boldsymbol{e}_y||\boldsymbol{n}|}=\frac{1}{|\boldsymbol{n}|},$$

$$\cos\widehat{\boldsymbol{e}_f\boldsymbol{n}}=\frac{\langle\boldsymbol{e}_f,\boldsymbol{n}\rangle}{|\boldsymbol{e}_f||\boldsymbol{n}|}=\frac{\frac{1}{p}x_0^2+\frac{p}{2}-\frac{1}{2p}x_0^2}{|\boldsymbol{n}|\sqrt{x_0^2+\left(\frac{p}{2}-\frac{1}{2p}x_0^2\right)^2}}$$

$$=\frac{\frac{p}{2}+\frac{1}{2p}x_0^2}{|\boldsymbol{n}|\sqrt{\left(\frac{p}{2}+\frac{1}{2p}x_0^2\right)^2}}=\frac{1}{|\boldsymbol{n}|}.$$

这样就证明了, 安放在抛物镜的焦点 $\left(0,\frac{p}{2}\right)$ 处的波源给出平行于镜轴 Oy 的光束, 而镜子使平行于 Oy 轴射来的光束会聚到焦点处 (见图 16).

例 7　我们用此例来表明, 切线只不过是函数图像在切点的邻域中的最佳线性近似, 它全然不必像在圆周的情形时那样, 或像在一般的凸曲线的情形时那样, 与函数图像只有唯一的公共点 (关于凸曲线将作专门论述.)

设函数 $f(x)$ 给定如下

$$f(x)=\begin{cases}x^2\sin\dfrac{1}{x}, & 若 x\neq 0,\\ 0, & 若 x=0.\end{cases}$$

这个函数的图像用粗线画在图 17 上.

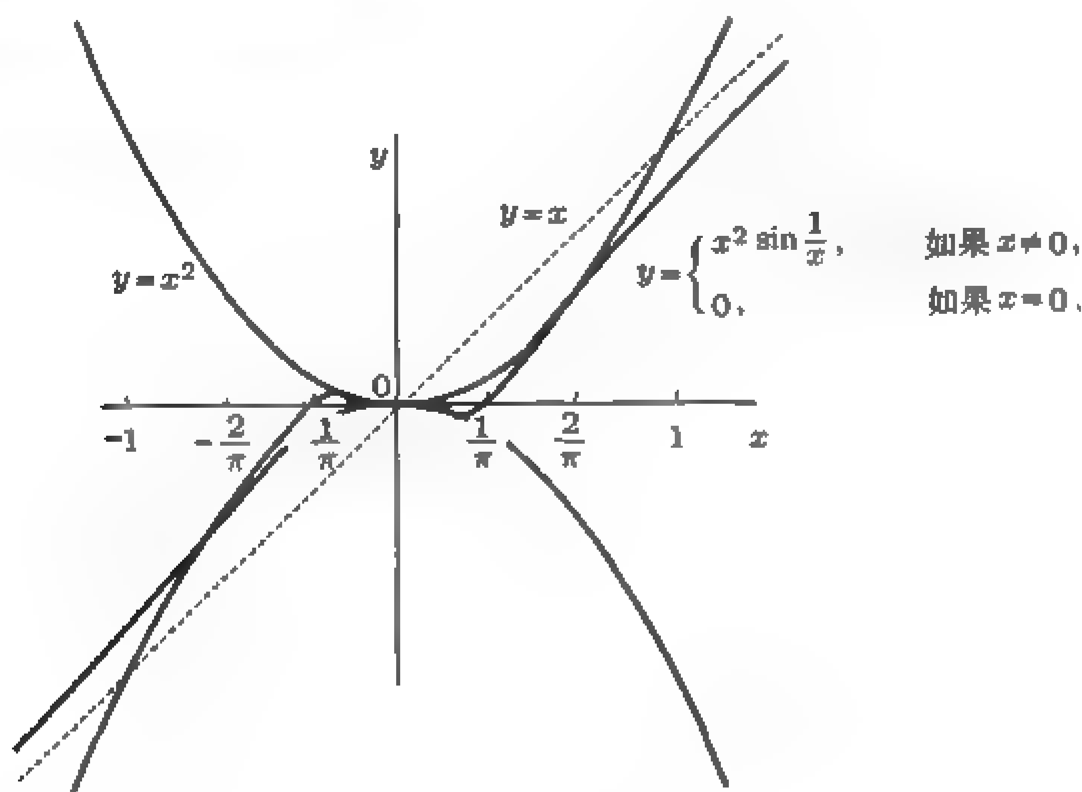


图 17

我们来求图像在点 (0,0) 处的切线. 由于

$$f'(0)=\lim_{x\to 0}\frac{x^2\sin\frac{1}{x}-0}{x-0}=\lim_{x\to 0}x\sin\frac{1}{x}=0,$$

所以, 切线的方程是 $y-0=0(x-0)$, 即 $y=0$.

于是, 在我们的例中, 切线与 Ox 轴重合, 图像与切线在切点的任一邻域中都有无穷多个交点.

根据函数 $f:E\to\mathbb{R}$ 在点 $x_0\in E$ 处可微的定义有

$$f(x)-f(x_0)=A(x_0)(x-x_0)+o(x-x_0),\ 当 x\to x_0, x\in E.$$

由于这个等式的右边当 $x\to x_0, x\in E$ 时趋于零, 那么,

$$\lim_{E\ni x\to x_0}f(x)=f(x_0),$$

因此, 在一点处可微的函数必定在此点处连续.

我们来证明反过来当然并不总是对的.

例 8 设 $f(x)=|x|$(图 18). 那么, 在点 $x_0=0$ 处

$$\lim_{x\to x_0-0}\frac{f(x)-f(x_0)}{x-x_0}=\lim_{x\to 0-0}\frac{|x|-0}{x-0}=\lim_{x\to 0-0}\frac{-x}{x}=-1,$$
$$\lim_{x\to x_0+0}\frac{f(x)-f(x_0)}{x-x_0}=\lim_{x\to 0+0}\frac{|x|-0}{x-0}=\lim_{x\to 0+0}\frac{x}{x}=1,$$

因此, 在这点处, 函数没有导数, 也就是说它在这点处不可微.

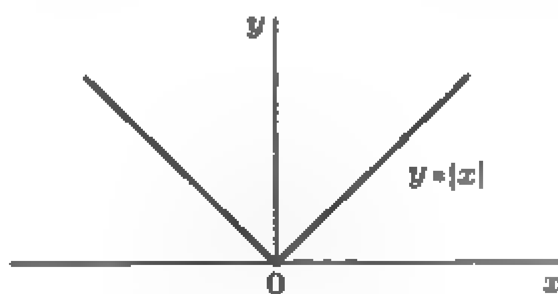


图 18

例 9 我们将证明: 当 $h\to 0$ 时, 有

$$e^{x+h}-e^x=e^xh+o(h).$$

因此, 函数 $\exp(x)=e^x$ 可微, 而且 $d\exp(x)h=\exp(x)h$, 即 $de^x=e^xdx$, 也就是说 $\exp' x=\exp x$, 即 $\dfrac{de^x}{dx}=e^x$.

◀ $e^{x+h}-e^x=e^x(e^h-1)=e^x(h+o(h))=e^xh+o(h)$.

我们使用了第三章 §2 第 4 段例 39 中得到的公式: 当 $h\to 0$ 时, $e^h-1=h+o(h)$. ▶

例 10 当 $h\to 0$ 且 $a>0$ 时,

$$a^{x+h}-a^x=a^x(\ln a)h+o(h).$$

于是 $da^x=a^x\ln a dx$, 从而, $\dfrac{da^x}{dx}=a^x\ln a$.

$$◀\ a^{x+h}-a^x=a^x(a^h-1)=a^x(e^{h\ln a}-1)$$
$$=a^x(h\ln a+o(h\ln a))=a^x(\ln a)h+o(h)\ \text{当}\ h\to 0.\ ▶$$

例 11　$\ln|x+h| - \ln|x| = \frac{1}{x}h + o(h)$,当$h \to 0$且$x \neq 0$. 于是, $d\ln|x| = \frac{1}{x}dx$, 从而, $\frac{d\ln|x|}{dx} = \frac{1}{x}$.

◀ $\ln|x+h| - \ln|x| = \ln\left|1+\frac{h}{x}\right|$.

当 $|h| < |x|$ 时, $\left|1+\frac{h}{x}\right| = 1+\frac{h}{x}$, 因此对于足够小的 h 值有

$$\ln|x+h| - \ln|x| = \ln\left(1+\frac{h}{x}\right) = \frac{h}{x} + o\left(\frac{h}{x}\right) = \frac{1}{x}h + o(h)$$

当 $h \to 0$. 此处我们用到了第三章 §2 第 4 段例 38 中所证明的公式: 当 $t \to 0$ 时, $\ln(1+t) = t + o(t)$. ▶

例 12　$\log_a|x+h| - \log_a|x| = \frac{1}{x\ln a}h + o(h)$, 当 $h \to 0$、$x \neq 0$、$0 < a \neq 1$ 时.

于是 $d\log_a|x| = \frac{1}{x\ln a}dx$, 从而, $\frac{d\log_a|x|}{dx} = \frac{1}{x\ln a}$.

$$◀ \log_a|x+h| - \log_a|x| = \log_a\left|1+\frac{h}{x}\right| = \log_a\left(1+\frac{h}{x}\right)$$
$$= \frac{1}{\ln a}\ln\left(1+\frac{h}{x}\right) = \frac{1}{\ln a}\left(\frac{h}{x} + o\left(\frac{h}{x}\right)\right) = \frac{1}{x\ln a}h + o(h).$$

我们使用了对数的换底公式和考察例 11 时的一些想法. ▶

练　习

1. 证明

a) 椭圆

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

在点 (x_0, y_0) 处的切线方程是

$$\frac{xx_0}{a^2} + \frac{yy_0}{b^2} = 1;$$

b) 由置于半轴为 $a > b > 0$ 的椭圆的两焦点 $F_1 = (-\sqrt{a^2-b^2}, 0)$、$F_2 = (\sqrt{a^2-b^2}, 0)$ 之一处的光源发出的光线被椭圆镜会聚在另一焦点处.

2. 写出数值的近似计算公式

a) $\sin\left(\frac{\pi}{6}+\alpha\right)$ 当 α 值接近于零时;

b) $\sin(30^\circ+\alpha^\circ)$ 当 α° 的值接近于零时;

c) $\cos\left(\frac{\pi}{4}+\alpha\right)$ 当 α 值接近于零时;

d) $\cos(45^\circ+\alpha^\circ)$ 当 α°的值接近于零时.

3. 盛着水的玻璃杯绕其轴以常角速度 ω 旋转, 设液体表面用通过旋转轴的平面来截所得的曲线的方程是 $y = f(x)$.

a) 证明 $f'(x)=\frac{\omega^2}{g}x$, 其中 g 是自由落体的加速度(见例 5).

b) 挑选满足 a) 中所述条件的函数 $f(x)$(见例 6).

c) 如果旋转的轴不与杯子的轴重合, 那么, 在 a) 中对函数所述的条件是否要改变?

4. 一个物体, 可以把它看作是一个质点, 在重力的作用下从一个光滑小山上滑下来山坡是可微函数 $y=f(x)$ 的图像.

a) 求物体在点 (x_0,y_0) 处的加速度向量的水平分量和竖直分量.

b) 当 $f(x)=x^2$, 且物体从很高的地方滑下来时, 求抛物线 $y=x^2$ 上使加速度的水平分量最大的点.

5. 处处不可导的连续周期函数例子, 置

$$\varPsi_0(x)=\begin{cases}x, & 当\quad 0\leqslant x\leqslant\frac{1}{2},\\ 1-x, & 当\quad \frac{1}{2}\leqslant x\leqslant 1,\end{cases}$$

并以周期 1 延拓这个函数到全数轴. 延拓后的周期函数用 φ_0 表示. 还设

$$\varphi_n(x)=\frac{1}{4^n}\varphi_0(4^n x).$$

函数 φ_n 周期为 4^{-n}, 且除了点 $x=\frac{k}{2^{n+1}}$ $(k\in\mathbb{Z})$ 以外处处有导数, 导数等于 1 或 -1. 设

$$f(x)=\sum_{n=1}^{\infty}\varphi_n(x).$$

证明, 函数 f 在 $\mathbb{R}$ 上有定义且连续, 但处处没有导数.(这个例子属于现代著名的荷兰数学家范德瓦尔登. 没有导数的连续函数的最初的例子是由波尔察诺 (1830 年) 和魏尔斯特拉斯 (1860 年) 构造的.)

§2. 微分的基本法则

构作已知函数的微分, 或等价地, 求它的导数, 叫做函数的微分运算①.

1. 微分法和算术运算

定理 1 *若函数 $f:X\to\mathbb{R}$、$g:X\to\mathbb{R}$ 都在点 $x\in X$ 处可微, 则*

a) 它们的和在 x 处可微, 且

$$(f+g)'(x)=(f'+g')(x);$$

b) 它们的积在 x 处可微, 且

$$(fg)'(x)=f'(x)g(x)+f(x)g'(x);$$

①虽说求微分的问题和求导数的问题是数学等价的, 可是导数和微分毕竟不是同一个事物, 因此, 例如说, 在法国的数学语言中就有两个术语: derivation —— “求导”, 求导数 (速度), 以及 differentiation —— “微分法”, 求微分.

c) 如果 $g(x) \neq 0$ 的话, 它们的比在 x 处可微, 且

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}.$$

◀ 我们用可微函数的定义和在第三章 §2 第 4 段建立的符号 $o(\ )$ 的性质来进行证明.

a) $$(f+g)(x+h)-(f+g)(x) = (f(x+h)+g(x+h)) - (f(x)+g(x))$$
$$= (f(x+h)-f(x)) + (g(x+h)-g(x))$$
$$= (f'(x)h + o(h)) + (g'(x)h + o(h))$$
$$= (f'(x)+g'(x))h + o(h)$$
$$= (f'+g')(x)h + o(h).$$

b) $$(f\cdot g)(x+h) - (f\cdot g)(x) = f(x+h)g(x+h) - f(x)g(x)$$
$$= (f(x+h)-f(x))g(x+h) + f(x)(g(x+h)-g(x))$$
$$= (f'(x)h+o(h))(g(x)+g'(x)h+o(h)) + f(x)(g'(x)h+o(h))$$
$$= f'(x)g(x)h + f(x)g'(x)h + o(h)$$
$$= (f'(x)g(x)+f(x)g'(x))h + o(h).$$

c) 由于在点 $x \in X$ 处可微的函数在该点连续, 所以, 注意到 $g(x) \neq 0$, 根据连续函数的性质, 即可断言, 对于足够小的 h 值, 同样有 $g(x+h) \neq 0$. 在下面的演算中假定 h 是足够小的:

$$\left(\frac{f}{g}\right)(x+h) - \left(\frac{f}{g}\right)(x) = \frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}$$
$$= \frac{1}{g(x)g(x+h)}(f(x+h)g(x) - f(x)g(x+h))$$
$$= \left(\frac{1}{g^2(x)} + o(1)\right)[(f(x)+f'(x)h+o(h))g(x) - f(x)(g(x)+g'(x)h+o(h))]$$
$$= \left(\frac{1}{g^2(x)} + o(1)\right)[(f'(x)g(x) - f(x)g'(x))h + o(h)]$$
$$= \frac{f'(x)g(x) - f(x)g'(x)}{g^2(x)}h + o(h).$$

我们用到了: 由函数 g 在点 x 的连续性以及 $g(x) \neq 0$ 得出的

$$\lim_{h\to 0}\frac{1}{g(x)g(x+h)} = \frac{1}{g^2(x)},$$

即

$$\frac{1}{g(x)g(x+h)} = \frac{1}{g^2(x)} + o(1),$$

这里, 当 $h \to 0$、$x+h \in X$ 时, $o(1)$ 是无穷小量. ▶

推论 1 可微函数的线性组合的导数等于这些函数的导数的线性组合.

◀ 因为常值函数显然是可微的, 且它的导数处处等于零, 所以只要在定理 1 的结论b) 中认为 $f \equiv \text{const} = c$ 就有 $(cg)'(x) = cg'(x)$.

现在用定理 1 的结论a), 可得

$$(c_1 f + c_2 g)'(x) = (c_1 f)'(x) + (c_2 g)'(x) = c_1 f'(x) + c_2 g'(x).$$

据此, 用归纳法便可证实

$$(c_1 f_1 + \cdots + c_n f_n)'(x) = c_1 f_1'(x) + \cdots + c_n f_n'(x). \ ▶$$

推论 2 *若函数 $f_1, \cdots, f_n$ 皆在点 x 处可微, 则*

$$\begin{aligned}(f_1 \cdot \cdots \cdot f_n)'(x) &= f_1'(x) f_2(x) \cdot \cdots \cdot f_n(x) \\ &+ f_1(x) f_2'(x) f_3(x) \cdot \cdots \cdot f_n(x) + \cdots \\ &+ f_1(x) \cdot \cdots \cdot f_{n-1}(x) f_n'(x).\end{aligned}$$

◀ 对于 $n = 1$ 结论是显然的.

若命题对于某 $n \in \mathbb{N}$ 成立, 则根据定理 1 的结论b), 它对于 $(n+1) \in \mathbb{N}$ 也成立. 根据归纳法原理, 断定所述公式对于任意的 $n \in \mathbb{N}$ 都成立. ▶

推论 3 *从导数和微分的相互关系推出, 定理 1 也可写成微分形式. 即:*

a) $d(f+g)(x) = df(x) + dg(x)$;

b) $d(f \cdot g)(x) = g(x)df(x) + f(x)dg(x)$;

c) $d\left(\dfrac{f}{g}\right)(x) = \dfrac{g(x)df(x) - f(x)dg(x)}{g^2(x)}$, 若 $g(x) \neq 0$.

我们来验证, 例如 a). 实际上,

$$\begin{aligned}d(f+g)(x)(h) &= (f+g)'(x)h = (f'+g')(x)h \\ &= (f'(x) + g'(x))h = f'(x)h + g'(x)h \\ &= df(x)(h) + dg(x)(h) = (df(x) + dg(x))(h),\end{aligned}$$

从而验证了 $d(f+g)(x)$ 和 $df(x) + dg(x)$ 是一样的.

例 1 *速度定义的不变性.* 现在我们来验证 §1 第 1 段中定义的质点的瞬时速度向量不依赖于笛卡儿坐标系的选择. 甚至对于任意一个仿射坐标系也能证实这一点.

设 (x^1, x^2) 和 $(\tilde{x}^1, \tilde{x}^2)$ 是平面的同一个点在两个不同的坐标系中的坐标, 这两个坐标系彼此之间的关系是

$$\begin{aligned}\tilde{x}^1 &= a_1^1 x^1 + a_2^1 x^2 + b^1, \\ \tilde{x}^2 &= a_1^2 x^1 + a_2^2 x^2 + b^2.\end{aligned} \tag{1}$$

因为任一向量 (在仿射空间中) 都由一对点决定, 而向量的坐标是它的终点和起点的坐标之差, 所以, 同一个向量在这两个坐标系中的坐标应以关系式

$$\begin{aligned}\tilde{v}^1 &= a_1^1 v^1 + a_2^1 v^2,\\ \tilde{v}^2 &= a_1^2 v^1 + a_2^2 v^2\end{aligned} \tag{2}$$

相联系.

如果点的运动规律在一个坐标系中由函数 $(x^1(t), x^2(t))$ 给出, 那么, 在另一个系中就由前面的关系式 (1) 相联系的函数 $(\tilde{x}^1(t), \tilde{x}^2(t))$ 来确定.

将关系式 (1) 关于时间 t 求微分, 按微分法则求出

$$\begin{aligned}\dot{\tilde{x}}^1 &= a_1^1 \dot{x}^1 + a_2^1 \dot{x}^2,\\ \dot{\tilde{x}}^2 &= a_1^2 \dot{x}^1 + a_2^2 \dot{x}^2.\end{aligned} \tag{3}$$

这样一来, 速度向量在第一个系中的坐标 $(v^1, v^2) = (\dot{x}^1, \dot{x}^2)$ 和它在第二个系中的坐标原来是以关系式 (2) 相联系的, 这就告诉我们, 我们遇到的是同一个向量的两种不同的写法.

例 2 设 $f(x) = \operatorname{tg} x$. 我们证明, 在 $\cos x \neq 0$ 处, 即在函数 $\operatorname{tg} x = \dfrac{\sin x}{\cos x}$ 的定义域内处处有 $f'(x) = \dfrac{1}{\cos^2 x}$.

在 §1 的例 1 和例 2 中曾证明,

$$\sin' x = \cos x, \quad \cos' x = -\sin x,$$

因此, 从定理 1 的结论 c) 得到, 当 $\cos x \neq 0$ 时, 有

$$\begin{aligned}\operatorname{tg}' x &= \left(\frac{\sin}{\cos}\right)'(x) = \frac{\sin' x \cos x - \sin x \cos' x}{\cos^2 x}\\ &= \frac{\cos x \cos x + \sin x \sin x}{\cos^2 x} = \frac{1}{\cos^2 x}.\end{aligned}$$

例 3 当 $\sin x \neq 0$ 时, 即在函数 $\operatorname{ctg} x = \dfrac{\cos x}{\sin x}$ 的定义域中,

$$\operatorname{ctg}' x = -\frac{1}{\sin^2 x}.$$

实际上,

$$\begin{aligned}\operatorname{ctg}' x &= \left(\frac{\cos}{\sin}\right)'(x) = \frac{\cos' x \sin x - \cos x \sin' x}{\sin^2 x}\\ &= -\frac{\sin x \sin x + \cos x \cos x}{\sin^2 x} = -\frac{1}{\sin^2 x}.\end{aligned}$$

例 4 若 $P(x) = c_0 + c_1 x + \cdots + c_n x^n$ 是多项式, 则 $P'(x) = c_1 + 2c_2 x + \cdots + nc_n x^{n-1}$.

实际上, 由于 $\dfrac{dx}{dx} = 1$, 那么, 根据推论 2, $\dfrac{dx^n}{dx} = nx^{n-1}$ 从而由推论 1 推出结论.

2. 复合函数的微分法

定理 2 (关于复合函数的微分) 如果函数 $f: X \to Y \subset \mathbb{R}$ 在点 $x \in X$ 处可微，而函数 $g: Y \to \mathbb{R}$ 在点 $y = f(x) \in Y$ 处可微，那么，这两个函数的复合 $g \circ f: X \to \mathbb{R}$ 在点 x 处可微，并且复合函数的微分

$$d(g \circ f)(x): T\mathbb{R}(x) \to T\mathbb{R}(g(f(x)))$$

等于两个微分

$$df(x): T\mathbb{R}(x) \to T\mathbb{R}(y = f(x)), \quad dg(y = f(x)): T\mathbb{R}(y) \to T\mathbb{R}(g(y))$$

的复合 $dg(y) \circ df(x)$.

◀ 函数 f 和 g 的可微性的条件表明

$$f(x+h) - f(x) = f'(x)h + o(h), \text{ 当} h \to 0, x+h \in X,$$
$$g(y+t) - g(y) = g'(y)t + o(t), \text{ 当} t \to 0, y+t \in Y.$$

我们看到，在后一个等式中，可以认为函数 $o(t)$ 在 $t = 0$ 也有定义，且有表达式 $o(t) = \gamma(t)t$，其中 $\gamma(0) = 0$，且当 $t \to 0, y+t \in Y$ 时，$\gamma(t) \to 0$. 令 $f(x) = y, f(x+h) = y + t$. 因为函数 f 在点 x 处可微，从而连续，我们断定，当 $h \to 0$ 时以及 $t \to 0$，若 $x + h \in X$，则 $y + t \in Y$. 根据关于复合函数的极限的定理，现在有

$$\gamma(f(x+h) - f(x)) = \alpha(h) \to 0, \text{ 当} h \to 0, x+h \in X \text{时},$$

于是，若 $t = f(x+h) - f(x)$，则

$$\begin{aligned} o(t) &= \gamma(f(x+h) - f(x))(f(x+h) - f(x)) \\ &= \alpha(h)(f'(x)h + o(h)) = \alpha(h)f'(x)h + \alpha(h)o(h) \\ &= o(h) + o(h) = o(h), \text{ 当} h \to 0, x+h \in X. \end{aligned}$$

还有

$$\begin{aligned} &(g \circ f)(x+h) - (g \circ f)(x) = g(f(x+h)) - g(f(x)) \\ &= g(y+t) - g(y) = g'(y)t + o(t) \\ &= g'(f(x))(f(x+h) - f(x)) + o(f(x+h) - f(x)) \\ &= g'(f(x))(f'(x)h + o(h)) + o(f(x+h) - f(x)) \\ &= g'(f(x))(f'(x)h) + g'(f(x))(o(h)) + o(f(x+h) - f(x)) \end{aligned}$$

由于量 $g'(f(x))(f'(x)h)$ 显然可以理解为映射 $h \overset{df(x)}{\longmapsto} f'(x)h$ 和映射 $\tau \overset{dg(y)}{\longmapsto} g'(y)\tau$ 的复合 $h \xmapsto{dg(y)\circ df(x)} g'(f(x)) \cdot f'(x)h$ 在平移 h 上的值 $(dg(f(x)) \circ df(x))(h)$. 因为我们已经确定

$$o(f(x+h)-f(x)) = o(h), 当h \to 0, x+h \in X,$$

为完成定理的证明, 只要再指出, 和式

$$g'(f(x))(o(h)) + o(f(x+h)-f(x))$$

与 h 相比, 当 $h \to 0$、$x+h \in X$ 时, 是无穷小量.

这样就证明了

$$\begin{aligned}&(g\circ f)(x+h)-(g\circ f)(x)\\ =\ &g'(f(x))f'(x)h+o(h), 当h \to 0、x+h \in X时. \blacktriangleright\end{aligned}$$

推论 4 *可微实值函数的复合的导数 $(g \circ f)'(x)$ 等于这些函数在相应的点处的导数的乘积 $g'(f(x)) \cdot f'(x)$.*

导数的富有内容的莱布尼茨符号是通往该命题的简捷证明的好响导. 按他的记法, 如果 $z = z(y), y = y(x)$, 就有

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx},$$

如果不把记号 $\frac{dz}{dy}$ 或 $\frac{dy}{dx}$ 看作是一个整体, 而分别把它们看作是 dz 与 dy 的比或 dy 与 dx 的比, 那么这个等式就是十分自然的了.

由此产生的证明思想是, 考虑差的比式

$$\frac{\Delta z}{\Delta x} = \frac{\Delta z}{\Delta y} \cdot \frac{\Delta y}{\Delta x},$$

然后让 $\Delta x \to 0$ 取极限. 这里发生的困难是, 即使 $\Delta x \neq 0, \Delta y$ 也可以是零 (这多少也是无法回避的困难!)

推论 5 *若有可微函数 $y_1 = f_1(x), \cdots, y_n = f_n(y_{n-1})$ 的复合 $(f_n \circ \cdots \circ f_1)(x)$, 则*

$$(f_n \circ \cdots \circ f_1)'(x) = f_n'(y_{n-1})f_{n-1}'(y_{n-2}) \cdot \cdots \cdot f_1'(x).$$

◀ 当 $n = 1$ 时结论是显然的.

若命题对于某 $n \in \mathbb{N}$ 成立, 则从定理 2 推出, 它对于 $n+1$ 也成立, 这就是说, 根据归纳法原理可确认它对于任意的 $n \in \mathbb{N}$ 成立. ▶

例 5 证明, 当 $\alpha \in \mathbb{R}$ 时, 在 $x > 0$ 的区域内有 $\frac{dx^\alpha}{dx} = \alpha x^{\alpha-1}$, 即 $dx^\alpha = \alpha x^{\alpha-1}dx$, 且 $(x+h)^\alpha - x^\alpha = \alpha x^{\alpha-1}h + o(h)$, 当 $h \to 0$ 时.

◀ 记 $x^\alpha = e^{\alpha \ln x}$ 使用刚证的定理, 并注意定理 1b) 款和 §1 中例 9、例 11 的那些结果.

设 $g(y) = e^y$ 及 $y = f(x) = \alpha \ln x$, 那么 $x^\alpha = (g \circ f)(x)$ 且

$$\begin{aligned}(g \circ f)'(x) &= g'(y)f'(x)\\ &= e^y \frac{\alpha}{x} = e^{\alpha \ln x}\frac{\alpha}{x} = x^\alpha \cdot \frac{\alpha}{x} = \alpha x^{\alpha-1}. \ \blacktriangleright\end{aligned}$$

例 6 可微函数模的对数的导数常常叫做对数导数.

由于 $F(x) = \ln|f(x)| = (\ln \circ |\ \ | \circ f)(x)$, 那么, 根据 §1 中例 11 的结果, $F'(x) = (\ln|f|)'(x) = \dfrac{f'(x)}{f(x)}$.

于是

$$d(\ln|f|)(x) = \frac{f'(x)}{f(x)}dx = \frac{df(x)}{f(x)}.$$

例 7 由自变量的取值误差引起的可微函数的值的绝对误差和相对误差.

若函数 f 在点 x 处可微, 则

$$f(x+h) - f(x) = f'(x)h + \alpha(x;h),$$

其中, 当 $h \to 0$ 时, $\alpha(x;h) = o(h)$.

于是, 如果在计算函数值 $f(x)$ 时, 自变量 x 的确定具有绝对误差 h, 那么, 由这个误差引起的函数值的绝对误差 $|f(x+h) - f(x)|$, 对于足够小的 h, 可以用在平移 h 上的微分的值的模 $|df(x)(h)| = |f'(x)h|$ 来代替.

此时, 可把比值

$$\frac{|f'(x)h|}{|f(x)|} = \frac{|df(x)(h)|}{|f(x)|}$$

或函数的对数导数与自变量的绝对误差的值的乘积的模 $\left|\frac{f'(x)}{f(x)}\right||h|$ 算作相对误差.

顺便指出, 如果 $f(x) = \ln x$, 那么, $d\ln x = \dfrac{dx}{x}$. 从而在确定对数值时的绝对误差等于在确定自变量值时的相对误差. 此事, 例如在对数尺上 (以及在其他许多具有非均匀刻度比例的仪器上) 得到了妙用. 那就是说, 我们想像数轴上零点右边的每个点都有自己的坐标 y, 把它标记在点的上方, 而在这点的下方标记数 $x = e^y$. 那么, $y = \ln x$. 同一个半轴被标上了一个均匀刻度 y 和一个不均匀刻度 x (这个不均匀刻度叫做对数刻度). 要想求出 $\ln x$, 应该把准线放在数 x 上并读出上面对应的数 y, 因为准线放置在某一点的准确性与对应于这个点的数 x 和 y 无关, 而放置准线的误差在均匀刻度中有某个值 Δy (可能偏差的线段的长度), 那么, 当按照数 x 来确定它的对数 y 时, 将大约有同一个绝对误差, 而当按照数的对数来确定这个数时, 将在一切刻度上大约都得到同一个相对误差.

例 8 我们来求函数 $u(x)^{v(x)}$ 的微分, 其中 $u(x)$ 和 $v(x)$ 都是可微函数, 并且 $u(x)>0$. 记 $u(x)^{v(x)}=e^{v(x)\ln u(x)}$, 并使用推论 5. 那么

$$\begin{aligned}&\frac{de^{v(x)\ln u(x)}}{dx}=e^{v(x)\ln u(x)}\left(v'(x)\ln u(x)+v(x)\frac{u'(x)}{u(x)}\right)\\&=u(x)^{v(x)}v'(x)\ln u(x)+v(x)u(x)^{v(x)-1}u'(x).\end{aligned}$$

3. 反函数的微分法

定理 3 (关于反函数的导数) *设函数 $f:X\to Y$ 和 $f^{-1}:Y\to X$ 互为反函数, 且分别在点 $x_0\in X$ 和 $f(x_0)=y_0\in Y$ 处连续. 如果函数 f 在点 x_0 处可微且 $f'(x_0)\neq 0$, 那么, 函数 f^{-1} 在点 y_0 也可微且*

$$(f^{-1})'(y_0)=(f'(x_0))^{-1}.$$

◀ 因为函数 $f:X\to Y$ 和 $f^{-1}:Y\to X$ 互为反函数, 所以, 当 $x\neq x_0, y=f(x)$ 时, $f(x)-f(x_0)$ 和 $f^{-1}(y)-f^{-1}(y_0)$ 都不为零. 从 f 在 x_0 的连续性和 f^{-1} 在 y_0 的连续性还可以断定, 当 $x\to x_0$ $x\in X$ 时, 有 $y=f(x)\to y_0, y=f(x)\in Y$, 并且, 若 $x\neq x_0$, 则 $y=f(x)\neq y_0$. 现在使用关于复合函数的极限的定理和极限的算术性质, 求出

$$\begin{aligned}\lim_{Y\ni y\to y_0}\frac{f^{-1}(y)-f^{-1}(y_0)}{y-y_0}&=\lim_{X\ni x\to x_0}\frac{x-x_0}{f(x)-f(x_0)}\\&=\lim_{X\ni x\to x_0}\frac{1}{\dfrac{f(x)-f(x_0)}{x-x_0}}=\frac{1}{f'(x_0)}.\end{aligned}$$

这就证明, 在点 y_0 处, 函数 $f^{-1}:Y\to X$ 有导数且

$$(f^{-1})'(y_0)=(f'(x_0))^{-1}. \blacktriangleright$$

注1 如果我们预先知道函数 f^{-1} 在点 y_0 处可微的话, 那么, 从恒等式 $(f^{-1}\circ f)(x)\equiv x$, 根据关于复合函数的微分法的定理就可立即求出 $(f^{-1})'(y_0)\cdot f'(x_0)=1$.

注2 条件 $f'(x_0)\neq 0$ 显然等价于: 由微分

$$df(x_0):T\mathbb{R}(x_0)\to T\mathbb{R}(y_0)$$

实现的映射 $h\mapsto f'(x_0)\,h$ 具有由公式 $\tau\mapsto(f'(x_0))^{-1}\tau$ 确定的逆映射

$$[df(x_0)]^{-1}:T\mathbb{R}(y_0)\to T\mathbb{R}(x_0).$$

这就是说, 使用微分的术语, 定理 3 的叙述中的第二句话可写成下面的形式:

如果函数 f 在点 x_0 处可微且在这点处它的微分 $df(x_0): T\mathbb{R}(x_0) \to T\mathbb{R}(y_0)$ 可逆, 那么 f 的反函数 f^{-1} 在点 $y_0 = f(x_0)$ 处存在微分且此微分就是映射 $df(x_0): T\mathbb{R}(x_0) \to T\mathbb{R}(y_0)$ 的逆映射

$$df^{-1}(y_0) = [df(x_0)]^{-1} : T\mathbb{R}(y_0) \to T\mathbb{R}(x_0).$$

例 9 我们证明, 当 $|y| < 1$ 时, $\arcsin' y = \dfrac{1}{\sqrt{1-y^2}}$.

函数 $\sin : \left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right] \to [-1,1]$ 和 $\arcsin : [-1,1] \to \left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right]$ 互为反函数且连续 (见第四章, §2, 例 8). 如果 $|x| < \dfrac{\pi}{2}$, 则 $\sin' x = \cos x \neq 0$, 且对于值 $y = \sin x$ 有 $|y| < 1$.

于是按定理 3

$$\arcsin' y = \frac{1}{\sin' x} = \frac{1}{\cos x} = \frac{1}{\sqrt{1-\sin^2 x}} = \frac{1}{\sqrt{1-y^2}}.$$

根式前面的符号是根据当 $|x| < \dfrac{\pi}{2}$ 时, $\cos x > 0$ 来选取的.

例 10 像上例中那样进行讨论可以证明 (注意第四章, §2 的例 9)

$$\arccos' y = -\frac{1}{\sqrt{1-y^2}}, \text{当} |y| < 1 \text{时}.$$

实际上

$$\arccos' y = \frac{1}{\cos' x} = -\frac{1}{\sin x} = -\frac{1}{\sqrt{1-\cos^2 x}} = -\frac{1}{\sqrt{1-y^2}}.$$

根式前面的符号是根据当 $0 < x < \pi$ 时, $\sin x > 0$ 来选取的.

例 11 $\operatorname{arctg}' y = \dfrac{1}{1+y^2}, \quad y \in \mathbb{R}$.

实际上

$$\operatorname{arctg}' y = \frac{1}{\operatorname{tg}' x} = \frac{1}{\dfrac{1}{\cos^2 x}} = \cos^2 x = \frac{1}{1+\operatorname{tg}^2 x} = \frac{1}{1+y^2}.$$

例 12 $\operatorname{arcctg}' y = -\dfrac{1}{1+y^2}, \quad y \in \mathbb{R}$.

实际上

$$\operatorname{arcctg}' y = \frac{1}{\operatorname{ctg}' x} = \frac{1}{-\dfrac{1}{\sin^2 x}} = -\sin^2 x = -\frac{1}{1+\operatorname{ctg}^2 x} = -\frac{1}{1+y^2}.$$

例 13 我们已经知道 (见 §1 中的例 10 和例 12), 函数 $y=f(x)=a^x$ 和 $x=f^{-1}(y)=\log_a x$ 有导数 $f'(x)=a^x\ln a$ 和 $(f^{-1})'(y)=\dfrac{1}{y\ln a}$.

我们来验证这符合定理 3:

$$(f^{-1})'(y)=\frac{1}{f'(x)}=\frac{1}{a^x\ln a}=\frac{1}{y\ln a},$$
$$f'(x)=\frac{1}{(f^{-1})'(y)}=\frac{1}{(y\ln a)^{-1}}=y\ln a=a^x\ln a.$$

例 14 *双曲、反双曲函数以及它们的导数.*

函数

$$\operatorname{sh} x=\frac{1}{2}(e^x-e^{-x}),$$
$$\operatorname{ch} x=\frac{1}{2}(e^x+e^{-x})$$

分别叫做 x 的*双曲正弦*和*双曲余弦*①.

以后会明白, 我们此刻纯形式地引入的这两个函数会像圆函数 $\sin x$、$\cos x$ 那样自然地出现在许多问题之中.

我们指出

$$\operatorname{sh}(-x)=-\operatorname{sh} x,$$
$$\operatorname{ch}(-x)=\operatorname{ch} x,$$

即, 双曲正弦是奇函数, 而双曲余弦是偶函数.

此外, 下列恒等式显然成立

$$\operatorname{ch}^2 x-\operatorname{sh}^2 x=1.$$

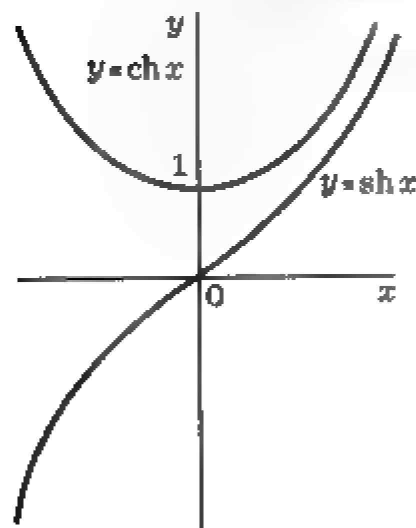


图 19

函数 $\operatorname{sh} x$ 和 $\operatorname{ch} x$ 的图像画在图 19 上.

从函数 $y=\operatorname{sh} x$ 的定义和函数 e^x 的性质推出, $\operatorname{sh} x$ 是一个连续的严格增的双方单值地映 $\mathbb{R}$ 到 $\mathbb{R}$ 的函数. 于是 $\operatorname{sh} x$ 有反函数, 它定义在 $\mathbb{R}$ 上, 连续并且严格单调增.

人们把此反函数记作

$$\operatorname{arsh} y$$

(读作 y 的*面积正弦*②).

①根据拉丁语 sinus hyperbolici, cosinus hyperbolici.

②全称是 area sinus hyperbolici; 此处为什么用面积 (area) 而不像在圆函数时那样用弧 (arcus), 以后再作解释.

这个函数容易通过已知的函数表示出来. 关于 x 解方程

$$\frac{1}{2}(e^x - e^{-x}) = y,$$

依次求得

$$e^x = y + \sqrt{1+y^2}$$

($e^x > 0$, 因此 $e^x \neq y - \sqrt{1+y^2}$) 和

$$x = \ln(y + \sqrt{1+y^2}).$$

于是

$$\operatorname{arsh} y = \ln(y + \sqrt{1+y^2}), \quad y \in \mathbb{R}.$$

类似地, 利用函数 $y = \operatorname{ch} x$ 在区间

$$\mathbb{R}_- = \{x \in \mathbb{R} | x \leqslant 0\}, \quad \mathbb{R}_+ = \{x \in \mathbb{R} | 0 \leqslant x\}$$

上的单调性, 可以建立函数 $\operatorname{arch}_- y$ 和 $\operatorname{arch}_+ y$, 它们对 $y \geqslant 1$ 定义, 分别是函数 $\operatorname{ch} x$ 在 $\mathbb{R}_-$ 和 $\mathbb{R}_+$ 上的限制的反函数.

它们由下面的公式给出

$$\operatorname{arch}_- y = \ln(y - \sqrt{y^2-1}),$$
$$\operatorname{arch}_+ y = \ln(y + \sqrt{y^2-1}).$$

从所引的定义求得

$$\operatorname{sh}' x = \frac{1}{2}(e^x + e^{-x}) = \operatorname{ch} x,$$
$$\operatorname{ch}' x = \frac{1}{2}(e^x - e^{-x}) = \operatorname{sh} x,$$

而根据关于反函数的导数的定理得到

$$\operatorname{arsh}' y = \frac{1}{\operatorname{sh}' x} = \frac{1}{\operatorname{ch} x} = \frac{1}{\sqrt{1+\operatorname{sh}^2 x}} = \frac{1}{\sqrt{1+y^2}},$$
$$\operatorname{arch}'_- y = \frac{1}{\operatorname{ch}' x} = \frac{1}{\operatorname{sh} x} = -\frac{1}{\sqrt{\operatorname{ch}^2 x - 1}} = -\frac{1}{\sqrt{y^2-1}}, \quad y > 1,$$
$$\operatorname{arch}'_+ y = \frac{1}{\operatorname{ch}' x} = \frac{1}{\operatorname{sh} x} = \frac{1}{\sqrt{\operatorname{ch}^2 x - 1}} = \frac{1}{\sqrt{y^2-1}}, \quad y > 1.$$

后三个关系式可以使用反双曲函数arsh y 和 arch y 的明显表达式来验证.

例如,

$$\operatorname{arsh}'y=\frac{1}{y+\sqrt{1+y^2}}\left(1+\frac{1}{2}\frac{1}{\sqrt{1+y^2}}2y\right)$$
$$=\frac{1}{y+\sqrt{1+y^2}}\cdot\frac{\sqrt{1+y^2}+y}{\sqrt{1+y^2}}=\frac{1}{\sqrt{1+y^2}}.$$

类似于 $\operatorname{tg}x$ 和 $\operatorname{ctg}x$ 可以考察函数

$$\operatorname{th}x=\frac{\operatorname{sh}x}{\operatorname{ch}x}\quad 和\quad \operatorname{cth}x=\frac{\operatorname{ch}x}{\operatorname{sh}x},$$

它们分别叫做*双曲正切*和*双曲余切*. 同样也可以考察它们的反函数*面积正切*

$$\operatorname{arth}y=\frac{1}{2}\ln\frac{1+y}{1-y},\quad |y|<1,$$

和面积余切:

$$\operatorname{arcth}y=\frac{1}{2}\ln\frac{y+1}{y-1},\quad |y|>1.$$

我们略去那些导出这些公式的求解初等方程的过程.

根据微分法则有

$$\operatorname{th}'x=\frac{\operatorname{sh}'x\operatorname{ch}x-\operatorname{sh}x\operatorname{ch}'x}{\operatorname{ch}^2x}=\frac{\operatorname{ch}x\operatorname{ch}x-\operatorname{sh}x\operatorname{sh}x}{\operatorname{ch}^2x}=\frac{1}{\operatorname{ch}^2x},$$
$$\operatorname{cth}'x=\frac{\operatorname{ch}'x\operatorname{sh}x-\operatorname{ch}x\operatorname{sh}'x}{\operatorname{sh}^2x}=\frac{\operatorname{sh}x\operatorname{sh}x-\operatorname{ch}x\operatorname{ch}x}{\operatorname{sh}^2x}=-\frac{1}{\operatorname{sh}^2x}.$$

根据关于反函数的导数的定理,

$$\operatorname{arth}'y=\frac{1}{\operatorname{th}'x}=\frac{1}{\frac{1}{\operatorname{ch}^2x}}=\operatorname{ch}^2x$$
$$=\frac{1}{1-\operatorname{th}^2x}=\frac{1}{1-y^2},\quad |y|<1,$$
$$\operatorname{arcth}'y=\frac{1}{\operatorname{cth}'y}=\frac{1}{-\frac{1}{\operatorname{sh}^2x}}=-\operatorname{sh}^2x$$
$$=-\frac{1}{\operatorname{cth}^2x-1}=-\frac{1}{y^2-1},\quad |y|>1.$$

这两个公式也可以通过直接对函数 $\operatorname{arth}y$ 和 $\operatorname{arcth}y$ 的明显表达式求微分来验证.

4. 基本初等函数的导数表　现在我们录下在 §1 和 §2 中算出的基本初等函数的导数 (见表 5.1).

表 5.1

函数 $f(x)$	导数 $f'(x)$	对自变量 $x \in \mathbb{R}$ 的变化区域的限制
1. C(常数)	0	
2. x^α	$\alpha x^{\alpha-1}$	当 $\alpha \in \mathbb{R}$ 时 $x > 0$
3. a^x	$a^x \ln a$	当 $\alpha \in \mathbb{N}$ 时 $x \in \mathbb{R}$
4. $\log_a \|x\|$	$\dfrac{1}{x \ln a}$	$x \in \mathbb{R}(a > 0, a \neq 1)$
5. $\sin x$	$\cos x$	$x \in \mathbb{R} \setminus 0(a > 0, a \neq 1)$
6. $\cos x$	$-\sin x$	
7. $\operatorname{tg} x$	$\dfrac{1}{\cos^2 x}$	$x \neq \dfrac{\pi}{2} + k\pi, k \in \mathbb{Z}$
8. $\operatorname{ctg} x$	$-\dfrac{1}{\sin^2 x}$	$x \neq k\pi, k \in \mathbb{Z}$
9. $\arcsin x$	$\dfrac{1}{\sqrt{1-x^2}}$	$\|x\| < 1$
10. $\arccos x$	$-\dfrac{1}{\sqrt{1-x^2}}$	$\|x\| < 1$
11. $\operatorname{arctg} x$	$\dfrac{1}{1+x^2}$	
12. $\operatorname{arcctg} x$	$-\dfrac{1}{1+x^2}$	
13. $\operatorname{sh} x$	$\operatorname{ch} x$	
14. $\operatorname{ch} x$	$\operatorname{sh} x$	
15. $\operatorname{th} x$	$\dfrac{1}{\operatorname{ch}^2 x}$	
16. $\operatorname{cth} x$	$-\dfrac{1}{\operatorname{sh}^2 x}$	$x \neq 0$
17. $\operatorname{arsh} x = \ln(x + \sqrt{1+x^2})$	$\dfrac{1}{\sqrt{1+x^2}}$	
18. $\operatorname{arch} x = \ln(x \pm \sqrt{x^2-1})$	$\pm\dfrac{1}{\sqrt{x^2-1}}$	$\|x\| > 1$
19. $\operatorname{arth} x = \dfrac{1}{2}\ln\dfrac{1+x}{1-x}$	$\dfrac{1}{1-x^2}$	$\|x\| < 1$
20. $\operatorname{arcth} x = \dfrac{1}{2}\ln\dfrac{x+1}{x-1}$	$\dfrac{1}{1-x^2}$	$\|x\| > 1$

5. 最简单的隐函数的微分法 设 $y = y(t)$ 和 $x = x(t)$ 是定义在点 $t_0 \in \mathbb{R}$ 的邻域 $U(t_0)$ 中的可微函数. 假定函数 $x = x(t)$ 有反函数 $t = t(x)$, 定义在点 $x_0 = x(t_0)$ 的邻域 $V(x_0)$ 中. 那么依赖于 t 的量 $y = y(t)$ 也可以看作是依隐蔽形式依赖于 x 的函数, 因为 $y(t) = y(t(x))$. 称它为隐函数. 假定 $x'(t_0) \neq 0$, 我们来求这个函数关于 x 在 x_0 点处的导数. 使用关于复合函数的微分的定理和关于反函数的微分的定理, 得到

$$y'_x|_{x=x_0} = \left.\frac{dy(t(x))}{dx}\right|_{x=x_0} = \left.\frac{dy(t)}{dt}\right|_{t=t_0} \cdot \left.\frac{dt(x)}{dx}\right|_{x=x_0}$$

$$= \frac{\left.\dfrac{dy(t)}{dt}\right|_{t=t_0}}{\left.\dfrac{dx(t)}{dt}\right|_{t=t_0}} = \frac{y'_t(t_0)}{x'_t(t_0)}.$$

(此处, 使用了通用的记号 $f(x)|_{x=x_0} := f(x_0)$).

如果同一个量看作是不同的自变量的函数, 那么, 为了避免误解, 在求微分时就需明确标明是对哪个变量求微分, 我们正是这样做的.

例 15　*速度的加法定律*　如果在给定计时系统的每个时刻 t 都知道点在直线上选定的坐标系 (数轴) 中的坐标 x, 那么, 点沿直线的运动就完全确定了. 于是, 数对 (x,t) 确定了点在时间推移过程中所处的空间位置. 点的运动规律可写成函数 $x=x(t)$ 的形式.

假设我们希望用另一个坐标系 $(\tilde{x},\tilde{t})$ 来描述同一个点的运动. 例如, 新的数轴相对于头一个数轴以速度 $-v$ 匀速运动 (在这种情况下可以把速度向量等同于确定它的那个数). 为简单起见, 我们认为在两个坐标系中坐标 (0,0) 属于同一个点, 或者精确地说, 在时刻 $\tilde{t}=0$, 点 $\tilde{x}=0$ 恰与时钟指示 $t=0$ 时的点 $x=0$ 重合.

那么, 经典的伽利略变换

$$\begin{aligned}\tilde{x}&=x+vt,\\ \tilde{t}&=t.\end{aligned}\tag{4}$$

就提供了联系描述在不同坐标系中观察同一个点的运动的坐标 (x,t) 和 $(\tilde{x},\tilde{t})$ 的一种可能的方式.

考察更一般的线性关系

$$\begin{aligned}\tilde{x}&=\alpha x+\beta t,\\ \tilde{t}&=\gamma x+\delta t,\end{aligned}\tag{5}$$

当然, 作为假定, 这个关系是可逆的, 即矩阵 $\begin{pmatrix}\alpha & \beta\\ \gamma & \delta\end{pmatrix}$ 的行列式异于零.

设 $x=x(t)$ 和 $\tilde{x}=\tilde{x}(\tilde{t})$ 是在这两个坐标系中所观察的点的运动规律.

如若知道了关系 $x=x(t)$, 从公式 (5) 就能求得

$$\begin{aligned}\tilde{x}(t)&=\alpha x(t)+\beta t,\\ \tilde{t}(t)&=\gamma x(t)+\delta t,\end{aligned}\tag{6}$$

而根据变换 (5) 的可逆性写下

$$\begin{aligned}x&=\tilde{\alpha}\tilde{x}+\tilde{\beta}\tilde{t},\\ t&=\tilde{\gamma}\tilde{x}+\tilde{\delta}\tilde{t},\end{aligned}\tag{7}$$

如若知道 $\tilde{x}=\tilde{x}(\tilde{t})$ 就可求出

$$\begin{aligned}x(\tilde{t})&=\tilde{\alpha}\tilde{x}(\tilde{t})+\tilde{\beta}\tilde{t},\\ t(\tilde{t})&=\tilde{\gamma}\tilde{x}(\tilde{t})+\tilde{\delta}\tilde{t}.\end{aligned}\tag{8}$$

从关系式 (6) 和 (8) 见到, 对于给定的点存在着互逆的关系 $\tilde{t}=\tilde{t}(t)$ 和 $t=t(\tilde{t})$.

现在来考察我们的点分别在坐标系 (x,t) 和 $(\tilde{x},\tilde{t})$ 中的速度

$$V(t)=\frac{dx(t)}{dt}=\dot{x}_t(t) \text{ 和 } \tilde{V}(\tilde{t})=\frac{d\tilde{x}(\tilde{t})}{d\tilde{t}}=\dot{\tilde{x}}_{\tilde{t}}(\tilde{t})$$

之间的联系.

使用隐函数的微分法则及公式 (6), 有

$$\frac{d\tilde{x}}{d\tilde{t}}=\frac{\dfrac{d\tilde{x}}{dt}}{\dfrac{d\tilde{t}}{dt}}=\frac{\alpha\dfrac{dx}{dt}+\beta}{\gamma\dfrac{dx}{dt}+\delta},$$

即

$$\tilde{V}(\tilde{t})=\frac{\alpha V(t)+\beta}{\gamma V(t)+\delta}, \tag{9}$$

其中 t 和 $\tilde{t}$ 是同一个时刻在坐标系 (x,t) 和 $(\tilde{x},\tilde{t})$ 中的坐标. 这一点在使用公式 (9) 的简略写法

$$\tilde{V}=\frac{\alpha V+\beta}{\gamma V+\delta} \tag{10}$$

的情况永远都要注意.

在伽利略变换 (4) 的情形, 从 (10) 得到经典的速度加法定律

$$\tilde{V}=V+v. \tag{11}$$

由足够精确的试验确定了, 在真空中, 光总以确定的速度 c 传播而与发光物体的运动情形无关 (这已经成为狭义相对论的一条公设) . 这就表示, 如果在时刻 $t=\tilde{t}=0$ 在点 $x=\tilde{x}=0$ 处突然闪光, 那么, 经过时间 t 在坐标系 (x,t) 中光达到的点的坐标 $x^2=(ct)^2$, 而在坐标系 $(\tilde{x},\tilde{t})$ 中这个事件将对应于时间 $\tilde{t}$ 以及这个点的坐标 $\tilde{x}$, 它满足 $\tilde{x}^2=(c\tilde{t})^2$.

这样一来, 如果 $x^2-c^2t^2=0$, 那么, 也有 $\tilde{x}^2-c^2\tilde{t}^2=0$, 反之亦然. 根据某些补充的物理的理由, 应该认为, 一般地, 如果 (x,t) 和 $(\tilde{x},\tilde{t})$ 在以关系式 (5) 相联系的不同的坐标系中对应同一个事件的话, 则

$$x^2-c^2t^2=\tilde{x}^2-c^2\tilde{t}^2. \tag{12}$$

条件 (12) 给出了变换 (5) 的系数 $\alpha,\beta,\gamma,\delta$ 的以下关系式:

$$\begin{aligned}\alpha^2-c^2\gamma^2&=1,\\ \alpha\beta-c^2\gamma\delta&=0,\\ \beta^2-c^2\delta^2&=-c^2.\end{aligned} \tag{13}$$

倘若 $c-1$, 则代替 (13) 我们应该有

$$\begin{aligned}&\alpha^2-\gamma^2=1,\\&\frac{\beta}{\delta}=\frac{\gamma}{\alpha},\\&\beta^2-\delta^2=-1.\end{aligned}\tag{14}$$

由此容易推出, 组 (14) 的通解可以按下列形式给出 (数对 (α,β)、(γ,δ) 可以改变符号):

$$\alpha=\operatorname{ch}\varphi,\quad \gamma=\operatorname{sh}\varphi,\quad \beta=\operatorname{sh}\varphi,\quad \delta=\operatorname{ch}\varphi,$$

其中 φ 是某个参数.

于是, (13) 的通解有如下的形式

$$\begin{pmatrix}\alpha&\beta\\\gamma&\delta\end{pmatrix}=\begin{pmatrix}\operatorname{ch}\varphi&c\operatorname{sh}\varphi\\\frac{1}{c}\operatorname{sh}\varphi&\operatorname{ch}\varphi\end{pmatrix}.$$

从而, 变换 (5) 可具体写作:

$$\begin{aligned}&\tilde{x}=\operatorname{ch}\varphi x+c\operatorname{sh}\varphi\, t,\\&\tilde{t}=\frac{1}{c}\operatorname{sh}\varphi\, x+\operatorname{ch}\varphi\, t.\end{aligned}\tag{15}$$

这就是洛伦兹变换.

为了弄清楚怎样来确定自由参数 φ, 我们想起轴 $\tilde{x}$ 是以速度 $-v$ 相对于轴 x 运动的, 即这个轴的点 $\tilde{x}=0$ 在系统 (x,t) 中观察具有速度 $-v$. 在 (15) 中置 $\tilde{x}=0$ 就求出它在系统 (x,t) 中的运动规律:

$$x=-c\operatorname{th}\varphi\, t.$$

于是,

$$\operatorname{th}\varphi=\frac{v}{c}.\tag{16}$$

把速度变换的一般规律 (10) 和洛伦兹变换 (15) 相对比, 得到

$$\tilde{V}=\frac{\operatorname{ch}\varphi V+c\operatorname{sh}\varphi}{\frac{1}{c}\operatorname{sh}\varphi V+\operatorname{ch}\varphi},$$

考虑到 (16), 得

$$\tilde{V}=\frac{V+v}{1+\frac{vV}{c^2}}.\tag{17}$$

公式 (17) 就是相对论的速度加法定律. 当 $|vV|\ll c^2$, 即当 $c\to\infty$ 时, 它就转化为由公式 (11) 表出的经典的速度加法定律.

注意到关系式 (16), 洛伦兹变换本身可以写成下面更自然的形式:

$$\tilde{x} = \frac{x + vt}{\sqrt{1 - \left(\frac{v}{c}\right)^2}},$$

$$\tilde{t} = \frac{t + \frac{v}{c^2}x}{\sqrt{1 - \left(\frac{v}{c}\right)^2}}. \tag{18}$$

由此可见, 当 $|v| \ll c$ 时, 即当 $c \to \infty$ 时, 它就变成经典的伽利略变换 (4).

6. 高阶导数 如果函数 $f : E \to \mathbb{R}$ 在任何一点 $x \in E$ 都可微, 那么, 在集 E 上就产生了一个新的函数 $f' : E \to \mathbb{R}$, 它在点 $x \in E$ 处的值等于函数 f 在这点处的导数 $f'(x)$.

函数 $f' : E \to \mathbb{R}$ 本身可以在 E 上有导数 $(f')' : E \to \mathbb{R}$, 这个导数, 相对于初始的函数 f, 叫做 f 的二阶导数, 并以符号

$$f''(x)、\frac{d^2 f(x)}{dx^2}$$

之一表示, 而若想明确标出微分的变元, 则在第一种情况下还写作 $f''_{xx}(x)$.

定义 按照归纳法, 如果 f 的 $n-1$ 阶导数 $f^{(n-1)}(x)$ 已定义, 则n 阶导数由公式

$$f^{(n)}(x) = (f^{(n-1)})'(x)$$

来定义.

对于 n 阶导数, 使用记号

$$f^{(n)}(x), \frac{d^n f(x)}{dx^n}.$$

约定 $f^{(0)}(x) := f(x)$.

一切在 E 上具有 n 阶连续导数的函数 $f : E \to \mathbb{R}$ 的集合用记号 $C^{(n)}(E, \mathbb{R})$ 表示, 而当不致引起混淆时, 用更简单的记号 $C^{(n)}(E)$ 或 $C^n(E, \mathbb{R})$ 或 $C^n(E)$ 来表示.

特别地, 根据已做的约定, 有 $f^{(0)}(x) = f(x), C^{(0)}(E) = C(E)$.

作为例子, 先来看算出的几个函数的高阶导数

例	$f(x)$	$f'(x)$	$f''(x)$	$\cdots$	$f^{(n)}(x)$
16.	a^x	$a^x \ln a$	$a^x \ln^2 a$	$\cdots$	$a^x \ln^n a$
17.	e^x	e^x	e^x	$\cdots$	e^x
18.	$\sin x$	$\cos x$	$-\sin x$	$\cdots$	$\sin\left(x + n\frac{\pi}{2}\right)$
19.	$\cos x$	$-\sin x$	$-\cos x$	$\cdots$	$\cos\left(x + n\frac{\pi}{2}\right)$

20. $(1+x)^\alpha$ $\alpha(1+x)^{\alpha-1}$ $\alpha(\alpha-1)(1+x)^{\alpha-2}$ $\cdots$ $\alpha(\alpha-1)\cdot\cdots\cdot(\alpha-n+1)(1+x)^{\alpha-n}$

21. x^α $\alpha x^{\alpha-1}$ $\alpha(\alpha-1)x^{\alpha-2}$ $\cdots$ $\alpha(\alpha-1)\cdot\cdots\cdot(\alpha-n+1)x^{\alpha-n}$

22. $\log_a|x|$ $\frac{1}{\ln a}x^{-1}$ $\frac{-1}{\ln a}x^{-2}$ $\cdots$ $\frac{(-1)^{n-1}(n-1)!}{\ln a}x^{-n}$

23. $\ln|x|$ x^{-1} $-x^{-2}$ $\cdots$ $(-1)^{n-1}(n-1)!x^{-n}$

例 24 莱布尼茨公式 设 $u(x)$ 和 $v(x)$ 同在集 E 上有直到 n 阶导数. 那么, 对于它们的乘积的 n 阶导数, 成立着下述莱布尼茨公式:

$$(uv)^{(n)} = \sum_{m=0}^{n} C_n^m u^{(n-m)}v^{(m)}. \tag{19}$$

莱布尼茨公式很像牛顿二项式 公式而且真的和它有直接的联系.

◀ 当 $n=1$ 时, 公式 (19) 与已建立的乘积的微分法则一样.

如果函数 u、v 有直到 $n+1$ 阶导数, 那么, 在公式 (19) 对于阶数 n 成立的假定下, 对它的左边和右边求微分之后得到

$$\begin{aligned}(uv)^{(n+1)} &= \sum_{m=0}^{n} C_n^m u^{(n-m+1)}v^{(m)} + \sum_{m=0}^{n} C_n^m u^{(n-m)}v^{(m+1)}\\ &= u^{(n+1)}v^{(0)} + \sum_{k=1}^{n}(C_n^k + C_n^{k-1})u^{(n+1-k)}v^{(k)} + u^{(0)}v^{(n+1)}\\ &= \sum_{k=0}^{n+1} C_{n+1}^k u^{(n+1-k)}v^{(k)}.\end{aligned}$$

我们合并了函数 u、v 的导数的乘积的同类项, 并且使用了公式 $C_n^k + C_n^{k-1} = C_{n+1}^k$.

这样一来, 用归纳法确定了莱布尼茨公式的真确性. ▶

例 25 若 $P_n(x) = c_0 + c_1x + \cdots + c_nx^n$, 则

$P_n(0) = c_0$

$P_n'(x) = c_1 + 2c_2x + \cdots + nc_nx^{n-1}$ 且 $P_n'(0) = c_1$,

$P_n''(x) = 2c_2 + 3\times 2c_3x + \cdots + n(n-1)c_nx^{n-2}$ 且 $P_n''(0) = 2!c_2$,

$P_n^{(3)}(x) = 3\times 2c_3 + \cdots + n(n-1)(n-2)c_nx^{n-3}$ 且 $P_n^{(3)}(0) = 3!c_3$,

....

$P_n^{(n)}(x) = n(n-1)(n-2)\cdot\cdots\cdot 2c_n$ 且 $P_n^{(n)}(0) = n!c_n$,

$P_n^{(k)}(x) \equiv 0$ 当 $k > n$.

于是, 多项式可以写成这样

$$P_n(x) = P_n^{(0)}(0) + \frac{1}{1!}P_n^{(1)}(0)x + \frac{1}{2!}P_n^{(2)}(0)x^2 + \cdots + \frac{1}{n!}P_n^{(n)}(0)x^n.$$

例 26 利用莱布尼茨公式以及多项式的导数当求导的阶数高于多项式的阶数时恒等于零这一事实, 可以求出例如函数 $f(x)=x^2\sin x$ 的 n 阶导数:

$$\begin{aligned}f^{(n)}(x)&=\sin^{(n)}x\cdot x^2+C_n^1\sin^{(n-1)}x\cdot 2x+C_n^2\sin^{(n-2)}x\cdot 2\\&=x^2\sin\left(x+n\frac{\pi}{2}\right)+2nx\sin\left(x+(n-1)\frac{\pi}{2}\right)\\&\quad+\left[-n(n-1)\sin\left(x+n\frac{\pi}{2}\right)\right]\\&=(x^2-n(n-1))\sin\left(x+n\frac{\pi}{2}\right)-2nx\cos\left(x+n\frac{\pi}{2}\right).\end{aligned}$$

例 27 设 $f(x)=\operatorname{arctg}x$. 我们来求 $f^{(n)}(0)(n=0,1,2,\cdots)$ 的值.

由于 $f'(x)=\dfrac{1}{1+x^2}$, 所以 $(1+x^2)f'(x)=1$.

对后一等式应用莱布尼茨公式, 就得递推公式

$$(1+x^2)f^{(n+1)}(x)+2nxf^{(n)}(x)+n(n-1)f^{(n-1)}(x)=0,$$

由此即可逐次求出函数 $f(x)$ 的一切导数.

置 $x=0$, 得到

$$f^{(n+1)}(0)=-n(n-1)f^{(n-1)}(0).$$

当 $n=1$ 时, 有 $f^{(2)}(0)=0$, 因此, 总有 $f^{(2m)}(0)=0$. 对于任意的奇阶数有

$$f^{(2m+1)}(0)=-2m(2m-1)f^{(2m-1)}(0),$$

又由于 $f'(0)=1$, 得到

$$f^{(2m+1)}(0)=(-1)^m(2m)!.$$

例 28 *加速度*. 如果 $x=x(t)$ 是沿数轴运动的质点的坐标对时间的依赖关系, 那么, $\dfrac{dx(t)}{dt}=\dot{x}(t)$ 是点的速度, 而 $\dfrac{d\dot{x}(t)}{dt}=\dfrac{d^2x(t)}{dt^2}=\ddot{x}(t)$ 是点在时刻 t 的加速度.

如果 $x(t)=\alpha t+\beta$, 那么 $\dot{x}(t)=\alpha$ 而 $\ddot{x}(t)\equiv 0$, 也就是说在匀速运动中加速度等于零. 很快我们就会验证, 若函数的二阶导数等于零, 则函数本身有 $\alpha t+\beta$ 的形式. 这样一来, 在匀速运动中并且仅在匀速运动中加速度等于零.

但当我们从两个坐标系中来观察在真空中依惯性运动的物体时, 如果我们希望在每个坐标系中物体都做匀速直线运动, 那就必须使得从一个惯性系到另一个系的变换公式是线性的. 正是由于这个缘故, 在例 15 中选择了线性的坐标变换公式 (5).

例 29 *最简单的隐函数的二阶导数*. 设 $y=y(t)$ 和 $x=x(t)$ 都是二阶可微函数. 假设函数 $x=x(t)$ 有可微的反函数 $t=t(x)$, 从而可认为量 $y(t)$ 依隐避形式依赖于 x, 这是因为 $y=y(t)=y(t(x))$. 在 $x'(t)\neq 0$ 的假定下我们来求二阶导数 y''_{xx}.

按照在第 5 款中考察的那种函数的微分法则, 有

$$y'_x = \frac{y'_t}{x'_t},$$

因此

$$y''_{xx} = (y'_x)'_x = \frac{(y'_x)'_t}{x'_t}$$

$$= \frac{\left(\frac{y'_t}{x'_t}\right)'_t}{x'_t} = \frac{\frac{y''_{tt}x'_t - y'_t x''_{tt}}{(x'_t)^2}}{x'_t} = \frac{x'_t y''_{tt} - x''_{tt} y'_t}{(x'_t)^3}.$$

我们注意到, 这里出现的每个函数, 包括 y''_{xx} 在内, 它们的明显表达式都是依赖于 t 的. 但是, 对于具体的点 x, 在把 t 用对应于这个 x 值的 $t = t(x)$ 的值代换后, 就能求得 y''_{xx} 在点 x 的值

例如, 若 $y = e^t, x = \ln t$, 则

$$y'_x = \frac{y'_t}{x'_t} = \frac{e^t}{\frac{1}{t}} = te^t, \quad y''_{xx} = \frac{(y'_x)'_t}{x'_t} = \frac{e^t + te^t}{\frac{1}{t}} = t(t+1)e^t.$$

我们特意取了一个简单的例子, 在这个例中可以通过 x 明显地表示出 $t, t = e^x$, 且在 $y(t) = e^t$ 中代人 $t = e^x$ 之后, 就可求得对 x 的明显的依赖关系 $y = e^{e^x}$. 对最后这个函数求微分, 就可以检验上面所得结果的正确性.

显然, 依次应用公式

$$y^{(n)}_{x^n} = \frac{\left(y^{(n-1)}_{x^{n-1}}\right)'_t}{x'_t}$$

就可以求出任意阶的导数.

练　习

1. 设 $\alpha_0, \alpha_1, \cdots, \alpha_n$ 是给定的实数. 求出在固定的点 $x_0 \in \mathbb{R}$ 有导数

$$P_n^{(k)}(x_0) = \alpha_k \quad (k = 0, 1, \cdots, n)$$

的 n 阶多项式 $P_n(x)$.

2. 计算 $f'(x)$, 如果

a) $f(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right), & \text{当} x \neq 0, \\ 0, & \text{当} x = 0. \end{cases}$

b) $f(x) = \begin{cases} x^2 \sin\frac{1}{x}, & \text{当} x \neq 0, \\ 0, & \text{当} x = 0. \end{cases}$

c) 验证问题a) 中的函数在 $\mathbb{R}$ 上任意次可微, 且 $f^{(n)}(0) = 0$.

d) 证明问题b) 中的函数的导数在 $\mathbb{R}$ 上有定义, 但不是 $\mathbb{R}$ 上的连续函数.

e) 证明函数

$$f(x)=\begin{cases}\exp\left(-\dfrac{1}{(1+x)^2}-\dfrac{1}{(1-x)^2}\right), & \text{当} -1<x<1,\\ 0, & \text{当} |x|\geqslant 1\end{cases}$$

在 $\mathbb{R}$ 上任意次可微.

3. 设 $f\in C^{(\infty)}(\mathbb{R})$. 证明当 $x\neq 0$ 时

$$\frac{1}{x^{n+1}}f^{(n)}\left(\frac{1}{x}\right)=(-1)^n\frac{d^n}{dx^n}\left[x^{n-1}f\left(\frac{1}{x}\right)\right].$$

4. 设 f 是在 $\mathbb{R}$ 上可微的函数. 证明
   a) 若 f 是偶函数, 则 f' 是奇函数.
   b) 若 f 是奇函数, 则 f' 是偶函数.

5. 证明
   a) 函数 $f(x)$ 在点 x_0 可微当且仅当 $f(x)-f(x_0)=\varphi(x)(x-x_0)$, 其中 $\varphi(x)$ 是在 x_0 连续的函数 (并且在这种情况下 $\varphi(x_0)=f'(x_0)$).
   b) 若 $f(x)-f(x_0)=\varphi(x)(x-x_0)$ 且 $\varphi\in C^{(n-1)}(U(x_0))$, 其中 $U(x_0)$ 是点 x_0 的邻域, 则函数 $f(x)$ 在点 x_0 处有 n 阶导数 $f^{(n)}(x_0)$.

6. 举例说明, 定理 3 中 f^{-1} 在点 y_0 连续的条件不是多余的.

7. a) 质量分别为 m_1 和 m_2 的两个物体, 在仅有的相互吸引力作用下在空间运动. 试利用牛顿定律 (§ 1 的公式 (1) 和 (2)) 验证, 在运动过程中, 量

$$E=\left(\frac{1}{2}m_1v_1^2+\frac{1}{2}m_2v_2^2\right)+\left(-G\frac{m_1m_2}{r}\right)=:K+U$$

   不变, 这里 v_1 和 v_2 是物体的速度, r 是两物体之间的距离.
   b) 给出量 $E=K+U$ 以及它的各个组成部分的物理解释.
   c) 将结果推广到 n 个物体运动的情形.

§3. 微分学的基本定理

1. 费马引理和罗尔定理

定义 1 如果点 x_0 在集 E 中有一个邻域 $U_E(x_0)$ 使得函数 $f:E\to\mathbb{R}$ 在任一点 $x\in U_E(x_0)$ 处都满足 $f(x)\leqslant f(x_0)$(或者 $f(x)\geqslant f(x_0)$), 那么就称点 $x_0\in E\subset\mathbb{R}$ 为函数 f 的*局部极大*(或*极小*)*值点*, 而称函数在这点处的值为它的*局部极大*(或*极小*)*值*.

定义 2 如果在任意点 $x\in U_E(x_0)\backslash x_0=\mathring{U}_E(x_0)$ 处都成立严格的不等式 $f(x)<f(x_0)(f(x)>f(x_0))$, 那么点 $x_0\in E$ 称为函数 $f:E\to\mathbb{R}$ 的*严格局部极大*(*小*)*值点*, 而函数在这点处的值称为它的*严格局部极大* (*小*) *值*.

定义 3 局部极大值点和局部极小值点都叫做*局部极值点*, 而函数在这种点处的值叫做函数的*局部极值*.

例 1　设

$$f(x)=\begin{cases}x^2, \text{若} -1\leqslant x<2,\\ 4, \ \text{若} 2\leqslant x.\end{cases}$$

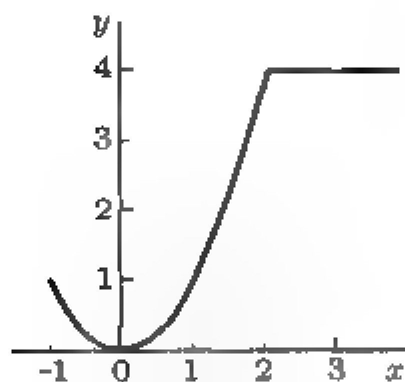


图 20

(图 20) 对于这个函数,

$x=-1$ 是严格局部极大值点;

$x=0$ 是严格局部极小值点;

$x=2$ 是局部极大值点;

$x>2$ 是极值点, 既是局部极大值点也是局部极小值点, 因为在这里函数局部为常数.

例 2　设 $f(x)=\sin\frac{1}{x}$ 在集合 $E=\mathbb{R}\backslash 0$ 上.

点 $x=\left(\frac{\pi}{2}+2k\pi\right)^{-1}, k\in\mathbb{Z}$ 是 $f(x)$ 的严格局部极大值点, 而点 $x=\left(-\frac{\pi}{2}+2k\pi\right)^{-1}, k\in\mathbb{Z}$ 是 $f(x)$ 的严格局部极小值点 (见图 12).

定义 4　函数 $f:E\to\mathbb{R}$ 的极值点 $x_0\in E$ 叫做内极值点, 如果 x_0 既是集 $E_-=\{x\in E|x<x_0\}$ 的极限点也是集 $E_+=\{x\in E|x>x_0\}$ 的极限点.

在例 2 中所有的极值点都是内极值点, 而在例 1 中点 $x=-1$ 不是内极值点.

引理 1 (费马)　*如果函数 $f:E\to\mathbb{R}$ 在内极值点 $x_0\in E$ 处可微, 则它在这点处的导数等于零: $f'(x_0)=0$.*

◀ 根据函数在点 x_0 处可微的定义

$$f(x_0+h)-f(x_0)=f'(x_0)h+\alpha(x_0;h)h,$$

其中, 当 $h\to 0, x_0+h\in E$ 时, $\alpha(x_0;h)\to 0$.

把这个关系式改写成

$$f(x_0+h)-f(x_0)=[f'(x_0)+\alpha(x_0;h)]h. \tag{1}$$

由于 x_0 是极值点, 对于一切足够接近于零且使 $x_0+h\in E$ 的 h 的值, 等式 (1) 的左边或者同时是非负的或者同时是非正的.

倘若 $f'(x_0)\neq 0$, 那么, 当 h 充分接近零时, 量 $f'(x_0)+\alpha(x_0;h)$ 与 $f'(x_0)$ 同号, 这是因为当 $h\to 0$、$x_0+h\in E$ 时, $\alpha(x_0;h)\to 0$.

至于说 h 本身的值, 则它既可以是正的也可以是负的, 因为 x_0 是内极值点.

这样一来, 如果 $f'(x_0)\neq 0$, 我们就得到这样的情形: 当 h 的符号变化时 (如果 h 足够接近于零), (1) 式的右边要改变符号, 同时 (1) 式的左边不能变号 (如果 h 足够接近于零). 这个矛盾就完成了证明. ▶

对费马引理的注 1° 费马引理给出了可微函数的内极值点的必要条件. 对于非内极值点 (像例 1 中的点 $x=-1$), 结论 $f'(x_0)=0$ 一般说来是不成立的.

2° 引理在几何上是十分明显的, 因为它断定在可微函数的极值点处, 函数图像的切线是水平的 (须知 $f'(x_0)$ 是切线对 Ox 轴的倾角的正切).

3° 引理在物理上表示, 沿直线运动的物体, 在开始返回时 (极值!) 的瞬时速度等于零.

从刚才证明的引理和关于闭区间上连续函数的最大 (小) 值的定理, 推出下面的

命题 1 (罗尔[①]定理) *若函数$f:[a,b]\to\mathbb{R}$在闭区间 $[a,b]$ 上连续, 在开区间 $]a,b[$ 中可微, 并且 $f(a)=f(b)$, 则存在点 $\xi\in]a,b[$, 使得$f'(\xi)=0$.*

◀ 因为函数 f 在闭区间 $[a,b]$ 上连续, 所以, 存在点 x_m、$x_M\in[a,b]$, 在这两点函数分别取它在这个闭区间上的最小值和最大值. 如果 $f(x_m)=f(x_M)$, 那么函数在 $[a,b]$ 上是常数, 而此时 $f'(x)\equiv 0$, 故结论显然成立. 若 $f(x_m)<f(x_M)$, 则因 $f(a)=f(b)$, 点 x_m 和 x_M 中必有一个落在开区间 $]a,b[$ 中, 我们把它记为 ξ. 根据费马引理, $f'(\xi)=0$. ▶

2. 拉格朗日和柯西的有限增量定理 下述命题是研究数值函数的一个最常用的重要工具.

定理 1 (关于有限增量的拉格朗日定理) *如果函数 $f:[a,b]\to\mathbb{R}$ 在闭区间 $[a,b]$ 上连续, 且在开区间 $]a,b[$ 中可微, 那么, 存在点 $\xi\in]a,b[$, 使得*

$$\boxed{f(b)-f(a)=f'(\xi)(b-a).} \tag{2}$$

◀ 为证定理, 考察辅助函数 $F(x)=f(x)-\dfrac{f(b)-f(a)}{b-a}(x-a)$. 它显然在闭区间 $[a,b]$ 上连续, 在开区间 $]a,b[$ 中可微, 且在它的端点处取相等的值 $F(a)=F(b)=f(a)$. 对 $F(x)$ 用罗尔定理, 得到点 $\xi\in]a,b[$, 使

$$F'(\xi)=f'(\xi)-\frac{f(b)-f(a)}{b-a}=0. \ ▶$$

对拉格朗日定理的注 1° 拉格朗日定理在几何上表示 (图 21) 在某一点 $(\xi,f(\xi))$ 处, 其中 $\xi\in]a,b[$, 函数图像的切线平行于联结点 $(a,f(a))$、$(b,f(b))$ 的弦, 因为后者的斜率等于 $\dfrac{f(b)-f(a)}{b-a}$.

---

①罗尔 (M. Rolle) (1652—1719)—— 法国数学家.

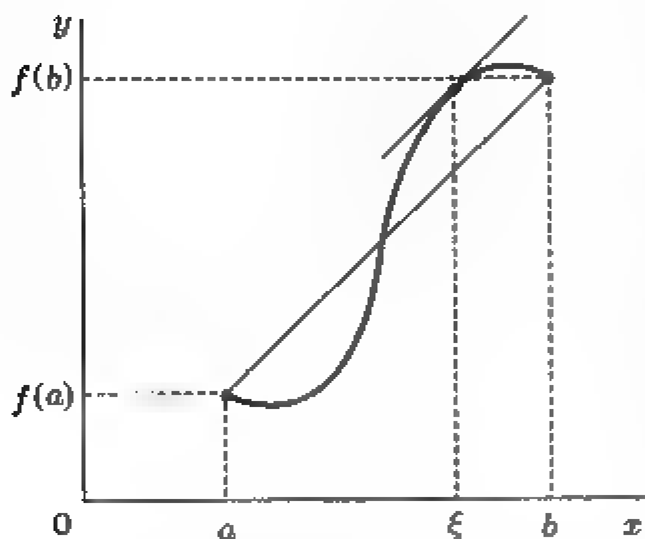


图 21

2° 若把 x 看作时间, 而把 $f(b)-f(a)$ 看作沿直线运动的质点在时间 $b-a$ 内的位移 (注意 $x \overset{f}{\longmapsto} f(x) \in \mathbb{R}$), 则拉格朗日定理表示, 质点的速度 $f'(x)$ 在某个时刻 $\xi \in]a,b[$ 具有这样的特点, 倘在整个时间进程 $[a,b]$ 中质点以常速度 $f'(\xi)$ 运动的话, 它就会得到同一个位移的值 $f(b)-f(a)$. 量 $f'(\xi)$ 自然被当作是在时间间隔 $[a,b]$ 中运动的平均速度.

3° 但须指出, 当运动不是沿直线进行的时候, 注 2° 意义下的平均速度可能不存在. 实际上, 例如设质点沿半径为 1 的圆周以常角速度 $\omega=1$ 运动. 我们知道, 它的运动规律可以写成

$$\boldsymbol{r}(t)=(\cos t, \sin t).$$

那么

$$\dot{\boldsymbol{r}}(t)=\boldsymbol{v}(t)=(-\sin t, \cos t).$$

于是 $|\boldsymbol{v}|=\sqrt{\sin^2 t+\cos^2 t}=1$.

在时刻 $t=0$ 和 $t=2\pi$, 质点位于平面的同一点 $\boldsymbol{r}(0)=\boldsymbol{r}(2\pi)=(1,0)$ 处, 于是等式

$$\boldsymbol{r}(2\pi)-\boldsymbol{r}(0)=\boldsymbol{v}(\xi)(2\pi-0)$$

应该表明 $\boldsymbol{v}(\xi)=\mathbf{0}$, 但这是不可能的.

不过我们明白, 在某一段时间内发生的位移与运动的速度之间当然是有联系的. 这种联系是, 通过的路程不可能超过速度的最大值与路上所花的时间的乘积. 此话可写成下面更确切的形式:

$$\boxed{|\boldsymbol{r}(b)-\boldsymbol{r}(a)| \leqslant \sup_{t\in]a,b[} |\dot{\boldsymbol{r}}(t)||b-a|.} \tag{3}$$

在适当的时候将证明, 这个很自然的不等式实际上是永远成立的. 这个不等式也叫做关于有限增量的拉格朗日定理, 而只对于数值函数成立的公式 (2) 常常叫做拉格朗日中间值定理 (中间值一词在这种情况下既指速度值 $f'(\xi)$ 是中间值, 也指点 ξ 位于 a、b 之间).

4° 拉格朗日定理之所以重要, 在于它把函数在有限闭区间上的增量同函数在这个区间上的导数联系起来了. 在此之前我们还不曾有过这种关于有限增量的定理, 而只是通过在固定点处的导数或微分来刻画函数的局部 (无穷小的) 增量.

拉格朗日定理的推论

推论 1 (函数单调性检验法) *如果在开区间的每一点函数的导数都是非负的 (或总是正的), 那么, 函数在这个开区间上不减 (或递增).*

◀ 实际上, 若 x_1, x_2 是区间中的两点, 且 $x_1 < x_2$, 即 $x_2 - x_1 > 0$, 则根据公式 (2) 有

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1), \text{其中} x_1 < \xi < x_2,$$

于是等式左边的差与 $f'(\xi)$ 同号. ▶

当然, 可以对具有非正 (或负的) 导数的函数的非增 (或递减) 性质叙述类似的命题.

注 根据关于反函数的定理和推论 1, 特别地, 可以断定, 如果在某个区间 I 上函数 $f(x)$ 有正的或者负的导数的话, 那么, 函数 f 在 I 上连续、在 I 上单调、有反函数 f^{-1}, 且 f^{-1} 定义在区间 $I' = f(I)$ 上并在 I' 上可微.

推论 2 (常值函数判别法) *在闭区间 $[a,b]$ 上连续的函数在此区间上为常数当且仅当它的导数在闭区间 $[a,b]$(或者甚至只要开区间 $]a,b[$) 的任一点都等于零.*

◀ 只需证明, 当 $f'(x)$ 在 $]a,b[$ 上恒为零时, 对于任意的 $x_1, x_2 \in [a,b]$, 都有 $f(x_1) = f(x_2)$. 但此事从拉格朗日定理推出, 按此定理

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1) = 0,$$

这是由于 ξ 位于 x_1 和 x_2 之间从而 $\xi \in]a,b[$, 于是 $f'(\xi) = 0$. ▶

注 由此显然可以做出下面的结论: 若两个函数 $F_1(x)$、$F_2(x)$ 的导数 $F_1'(x)$、$F_2'(x)$ 在某区间上重合, 即 $F_1'(x) = F_2'(x)$, 则在此区间上差 $F_1(x) - F_2(x)$ 是常值函数.(我们将会看到, 这个结论对于积分学是很重要的.)

以下是拉格朗日定理的有益的推广, 它也基于罗尔定理.

命题 2 (柯西有限增量定理) *设 $x = x(t)$ 及 $y = y(t)$ 是在闭区间 $[\alpha, \beta]$ 上连续且在开区间 $]\alpha, \beta[$ 中可微的函数. 那么存在点 $\tau \in]\alpha, \beta[$, 使得*

$$x'(\tau)\left(y(\beta) - y(\alpha)\right) = y'(\tau)\left(x(\beta) - x(\alpha)\right).$$

如果对于任意 $t \in]\alpha, \beta[$ 有 $x'(t) \neq 0$, 则 $x(\alpha) \neq x(\beta)$ 且成立等式

$$\frac{y(\beta) - y(\alpha)}{x(\beta) - x(\alpha)} = \frac{y'(\tau)}{x'(\tau)}. \tag{4}$$

◀ 函数 $F(t) = x(t)\left(y(\beta) - y(\alpha)\right) - y(t)\left(x(\beta) - x(\alpha)\right)$ 在闭区间 $[\alpha, \beta]$ 上满足罗尔定理的条件, 因此, 存在点 $\tau \in]\alpha, \beta[$ 使 $F'(\tau) = 0$, 它等价于要证的等式, 为了从

它推出关系式 (4), 只要再注意, 如果在 $]\alpha,\beta[$ 上有 $x'(t)\neq 0$, 则仍根据罗尔定理得 $x(\alpha)\neq x(\beta)$. ▶

柯西定理的注 1° 如果把一对函数 $x(t), y(t)$ 看作质点运动的规律, 那么 $(x'(t), y'(t))$ 是它在时刻 t 的速度向量, 而 $(x(\beta)-x(\alpha), y(\beta)-y(\alpha))$ 是它在时间间隔 $[\alpha,\beta]$ 内的位移向量, 从而定理断定, 在某时刻 $\tau\in]\alpha,\beta[$, 这两个向量共线. 但是这个关于平面运动的事实, 同直线运动中关于平均速度的定理一样, 只不过是一种令人喜欢的特殊现象. 事实上, 请想象一下沿着螺旋线均匀上升的质点吧. 它的速度与铅垂线构成不变的非零的夹角, 可是, 位移向量却可能是竖直的 (绕一圈所作的位移).

2° 拉格朗日公式可从柯西公式得到, 如果在柯西公式中令 $x=x(t)=t, y(t)=y(x)=f(x), \alpha=a, \beta=b$.

3. 泰勒[①]公式 从目前已经讲过的微分学这部分内容可以产生这样一个正确的观念, 两个函数在某点处越是有更多的导数 (包括零阶导数) 相同, 它们在这点的邻域内就越近似. 我们大体上已经注意过函数在某点的邻域内用多项式

$$P_n(x)=P_n(x_0;x)=c_0+c_1(x-x_0)+\cdots+c_n(x-x_0)^n$$

来近似的问题, 现在我们再来研究这个问题. 我们知道 (见 §2 第 6 段的例 25), 代数多项式可以表成

$$P_n(x)=P_n(x_0)+\frac{P_n'(x_0)}{1!}(x-x_0)+\cdots+\frac{P_n^{(n)}(x_0)}{n!}(x-x_0)^n,$$

即 $c_k=\dfrac{P_n^{(k)}(x_0)}{k!}(k=0,1,\cdots,n)$. 此事容易直接验证.

于是, 如果我们给定了一个在点 x_0 处有 n 阶导数的函数 $f(x)$, 那么我们可以随即写出多项式

$$P_n(x_0;x)=P_n(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n, \tag{5}$$

它在点 x_0 处不超过 n 阶的导数与函数 $f(x)$ 在点 x_0 处的同阶导数相同.

定义 5 由关系式 (5) 给出的多项式叫做函数 $f(x)$ 在点 x_0 处的 n 阶泰勒多项式.

我们感兴趣的是量

$$f(x)-P_n(x_0;x)=r_n(x_0;x), \tag{6}$$

即多项式 $P_n(x)$ 与函数 $f(x)$ 的偏差. 此量常被称为泰勒公式

$$\boxed{f(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+r_n(x_0;x)} \tag{7}$$

① 泰勒 (Taylor)(1685—1731)—— 英国数学家.

的余式, 或更确切地称为 n阶余式或 n阶泰勒公式余项.

如果对于函数 $r_n(x_0;x)$, 除了它的定义 (6) 外毫无所知, 那么等式 (7) 本身当然是毫无用处的.

现在我们使用技巧性相当强的方法来求关于余项的信息. 更自然的办法将在积分学中给出.

定理 2 如果在以 x_0、x 为端点的闭区间上函数 f 连同它的前 n 阶导数连续, 而在这个区间的内点处它有 $n+1$ 阶导数, 那么对于任意一个在这个闭区间上连续且在它的内点处有异于零的导数的函数 φ, 都存在位于 x_0 和 x 之间的点 ξ 使得

$$\boxed{r_n(x_0;x)=\frac{\varphi(x)-\varphi(x_0)}{\varphi'(\xi)n!}f^{(n+1)}(\xi)(x-\xi)^n.} \tag{8}$$

◀ 在以 x_0、x 为端点的闭区间 I 上考察辅助函数

$$F(t)=f(x)-P_n(t;x), \tag{9}$$

它是自变量 t 的函数. 详细写出 $F(t)$ 的定义:

$$F(t)=f(x)-\left[f(t)+\frac{f'(t)}{1!}(x-t)+\cdots+\frac{f^{(n)}(t)}{n!}(x-t)^n\right]. \tag{10}$$

从函数 $F(t)$ 的定义和定理的条件看出, F 在闭区间 I 上连续且在它的内点处可微, 同时

$$\begin{aligned}F'(t)=-\Big[f'(t)&-\frac{f'(t)}{1!}+\frac{f''(t)}{1!}(x-t)\\&-\frac{f''(t)}{2!}\cdot 2(x-t)+\frac{f'''(t)}{2!}(x-t)^2-\cdots\\&+\frac{f^{(n+1)}(t)}{n!}(x-t)^n\Big]=-\frac{f^{(n+1)}(t)}{n!}(x-t)^n.\end{aligned}$$

对于闭区间 I 上的函数对 $F(t)$、$\varphi(t)$ 用柯西定理 (见关系式 (4)), 求得介于 x_0 和 x 之间的点 ξ, 在这点处, 有

$$\frac{F(x)-F(x_0)}{\varphi(x)-\varphi(x_0)}=\frac{F'(\xi)}{\varphi'(\xi)}.$$

把 $F'(\xi)$ 的表达式代进来, 并注意到从公式 (6)、(9) 和 (10) 推出的 $F(x)-F(x_0)=0-F(x_0)=-r_n(x_0;x)$, 就得到公式 (8). ▶

在 (8) 中置 $\varphi(t)=x-t$, 得到

推论 1 (余项的柯西公式).

$$\boxed{r_n(x_0;x)=\frac{1}{n!}f^{(n+1)}(\xi)(x-\xi)^n(x-x_0).} \tag{11}$$

在 (8) 中置 $\varphi(t)=(x-t)^{n+1}$, 得到特别优美的公式

推论 2 (余项的拉格朗日公式)

$$\boxed{r_n(x_0;x)=\frac{1}{(n+1)!}f^{(n+1)}(\xi)(x-x_0)^{n+1}.} \tag{12}$$

我们指出, 泰勒公式 (7) 当 $x_0=0$ 时常叫做麦克劳林①公式.

看一些例子.

例 3　对于函数 $f(x)=e^x$, 当 $x_0=0$ 时, 泰勒公式表示为

$$e^x=1+\frac{1}{1!}x+\frac{1}{2!}x^2+\cdots+\frac{1}{n!}x^n+r_n(0;x), \tag{13}$$

根据 (12) 可以认为

$$r_n(0;x)=\frac{1}{(n+1)!}e^{\xi}x^{n+1},$$

其中 $|\xi|<|x|$.

于是

$$|r_n(0;x)|=\frac{1}{(n+1)!}e^{\xi}|x|^{n+1}<\frac{|x|^{n+1}}{(n+1)!}e^{|x|}. \tag{14}$$

但对于任意固定的 $x\in\mathbb{R}$, 当 $n\to\infty$ 时, 我们已知 (见第三章 §1 第 36 段) 量 $\frac{|x|^{n+1}}{(n+1)!}$ 是趋于零的. 这就表明, 从估计式 (14) 及级数和的定义推出, 对于 $x\in\mathbb{R}$ 有

$$e^x=1+\frac{1}{1!}x+\frac{1}{2!}x^2+\cdots+\frac{1}{n!}x^n+\cdots. \tag{15}$$

例 4　类似地, 对于任意的 $a, 0<a, a\neq 1$, 我们得到函数 a^x 的展开式:

$$a^x=1+\frac{\ln a}{1!}x+\frac{\ln^2 a}{2!}x^2+\cdots+\frac{\ln^n a}{n!}x^n+\cdots.$$

例 5　设 $f(x)=\sin x$. 我们知道 (见 §2 第 6 段例 18) $f^{(n)}(x)=\sin\left(x+n\frac{\pi}{2}\right)$, $n\in\mathbb{N}$, 因此, 从拉格朗日公式 (12), 对于 $x_0=0$ 和任意的 $x\in\mathbb{R}$, 求得

$$r_n(0;x)=\frac{1}{(n+1)!}\sin\left(\xi+\frac{\pi}{2}(n+1)\right)\cdot x^{n+1}. \tag{16}$$

由此推出, 对于任意的固定的值 $x\in\mathbb{R}$, 当 $n\to\infty$ 时, 量 $r_n(0;x)$ 趋于零. 于是对于任意的 $x\in\mathbb{R}$, 展开式成立:

$$\sin x=x-\frac{1}{3!}x^3+\frac{1}{5!}x^5-\cdots+\frac{(-1)^n}{(2n+1)!}x^{2n+1}+\cdots. \tag{17}$$

①麦克劳林 (C.Maclaurin)(1698—1746)—— 英国数学家.

例 6 类似地，对于函数 $f(x) = \cos x$ 得到

$$r_n(0;x) = \frac{1}{(n+1)!}\cos\left(\xi + \frac{\pi}{2}(n+1)\right)\cdot x^{n+1} \tag{18}$$

和

$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots + \frac{(-1)^n}{(2n)!}x^{2n} + \cdots. \tag{19}$$

例 7 由于 $\mathrm{sh}'x = \mathrm{ch}\, x$，而 $\mathrm{ch}'x = \mathrm{sh}\, x$，所以，对于函数 $f(x) = \mathrm{sh}\, x$，当 $x_0 = 0$ 时，从公式 (12) 得到

$$r_n(0;x) = \frac{1}{(n+1)!}\varphi(\xi)x^{n+1},$$

其中当 n 为偶数时，$\varphi(\xi) = \mathrm{sh}\, \xi$，而当 n 为奇数时，$\varphi(\xi) = \mathrm{ch}\, \xi$. 因为 $|\xi| < |x|$，在任何情况下都有

$$|\varphi(\xi)| \leqslant \max\{|\mathrm{sh}\, x|, |\mathrm{ch}\, x|\}.$$

这就表明，对于任意固定的值 $x \in \mathbb{R}$，当 $n \to \infty$ 时，$r_n(0,x) \to 0$. 于是我们得到对任意的 $x \in \mathbb{R}$ 都成立的展开式

$$\mathrm{sh}\, x = x + \frac{1}{3!}x^3 + \frac{1}{5!}x^5 + \cdots + \frac{1}{(2n+1)!}x^{2n+1} + \cdots. \tag{20}$$

例 8 类似地得到对于任意的值 $x \in \mathbb{R}$ 都成立的展开式

$$\mathrm{ch}\, x = 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \cdots + \frac{1}{(2n)!}x^{2n} + \cdots. \tag{21}$$

例 9 对于函数 $f(x) = \ln(1+x)$，有

$$f^{(n)}(x) = \frac{(-1)^{n-1}(n-1)!}{(1+x)^n},$$

因此，当 $x_0 = 0$ 时，泰勒公式 (7) 对于这个函数成为

$$\ln(x+1) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots + \frac{(-1)^{n-1}}{n}x^n + r_n(0;x). \tag{22}$$

这次我们按柯西公式 (11) 来表示 $r_n(0;x)$:

$$r_n(0;x) = \frac{1}{n!}\frac{(-1)^n n!}{(1+\xi)^n}(x-\xi)^n x,$$

或

$$r_n(0;x) = (-1)^n x \cdot \left(\frac{x-\xi}{1+\xi}\right)^n, \tag{23}$$

其中点 ξ 在 0 和 x 之间.

如果 $|x|<1$, 那么, 从 ξ 在 0 和 x 之间这个条件推出

$$\left|\frac{x-\xi}{1+\xi}\right| = \frac{|x|-|\xi|}{|1+\xi|} \leqslant \frac{|x|-|\xi|}{1-|\xi|} = 1-\frac{1-|x|}{1-|\xi|} \leqslant 1-\frac{1-|x|}{1-0} = |x|. \tag{24}$$

于是当 $|x|<1$ 时

$$|r_n(0,x)| \leqslant |x|^{n+1}. \tag{25}$$

因此, 当 $|x|<1$ 时成立展开式

$$\ln(1+x) = x-\frac{1}{2}x^2+\frac{1}{3}x^3-\cdots+\frac{(-1)^{n-1}}{n}x^n+\cdots. \tag{26}$$

我们发现, 在闭区间 $|x|\leqslant 1$ 之外, (26) 右端的级数处处发散, 因为它的一般项当 $|x|>1$ 时不趋于零.

例 10 若 $f(x)=(1+x)^\alpha$, 其中 $\alpha\in\mathbb{R}$, 则

$$f^{(n)}(x) = \alpha(\alpha-1)\cdots(\alpha-n+1)(1+x)^{\alpha-n}.$$

因此, 当 $x_0=0$ 时, 这个函数的泰勒公式 (7) 有如下的形式:

$$(1+x)^\alpha = 1+\frac{\alpha}{1!}x+\frac{\alpha(\alpha-1)}{2!}x^2+\cdots+\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n+r_n(0;x). \tag{27}$$

使用柯西公式 (11), 求得

$$r_n(0;x) = \frac{\alpha(\alpha-1)\cdots(\alpha-n)}{n!}(1+\xi)^{\alpha-n}(x-\xi)^n x, \tag{28}$$

其中 ξ 位于 0 和 x 之间.

如果 $|x|<1$, 那么使用估计式 (24), 有

$$|r_n(0;x)| \leqslant \left|\alpha\left(1-\frac{\alpha}{1}\right)\cdots\left(1-\frac{\alpha}{n}\right)\right|(1+\xi)^\alpha|x|^{n+1}. \tag{29}$$

当 n 增加 1 时, 不等式 (29) 的右边乘以 $\left|\left(1-\frac{\alpha}{n+1}\right)x\right|$. 但是, 因为 $|x|<1$, 所以, 如果 $|x|<q<1$, 则不管 α 的值如何, 对于足够大的 n 将有 $\left|\left(1-\frac{\alpha}{n+1}\right)x\right|<q<1$.

由此推出, 对于任意的 $\alpha\in\mathbb{R}$ 和开区间 $|x|<1$ 中任意的 x, 当 $n\to\infty$ 时, $r_n(0;x)\to 0$; 因此在开区间 $|x|<1$ 上成立由牛顿得到的展开式 (牛顿二项式)

$$\begin{aligned}(1+x)^\alpha = 1+\frac{\alpha}{1!}x+\frac{\alpha(\alpha-1)}{2!}x^2+\cdots\\ +\frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n+\cdots.\end{aligned} \tag{30}$$

我们注意到, 从达朗贝尔比较检验法 (见第三章 §1 第 46 段) 可以看出, 当 $|x|>1$ 时, 只要 $\alpha\notin\mathbb{N}$, 级数 (30) 总是发散的.

现在我们单独讨论 $\alpha = n \in \mathbb{N}$ 的情形.

这时函数 $f(x) = (1+x)^\alpha = (1+x)^n$ 是 n 次多项式, 因此它的一切高于 n 阶的导数都是零. 于是, 我们从泰勒公式 (7), 以及例如拉格朗日公式 (12) 都得到等式:

$$(1+x)^n = 1 + \frac{n}{1!}x + \frac{n(n-1)}{2!}x^2 + \cdots + \frac{n(n-1)\cdots 1}{n!}x^n, \tag{31}$$

这正是我们在读中学时就知道的具自然数指数的牛顿二项式公式:

$$(1+x)^n = 1 + C_n^1 x + C_n^2 x^2 + \cdots + C_n^n x^n.$$

这样, 我们已经定义了泰勒公式 (7), 并且得到了泰勒公式余项的表达式 (8)、(11)、(12). 我们所得到的关系式 (14)、(16)、(18)、(25)、(29) 可以用来估计按泰勒公式计算一些重要的初等函数时的误差. 最后, 我们还得到了这些函数的幂级数展开式.

定义 6 如果函数 $f(x)$ 在点 x_0 处有任意阶导数, 那么级数

$$f(x_0) + \frac{1}{1!}f'(x_0)(x-x_0) + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n + \cdots$$

叫做函数 f 在点 x_0 处的*泰勒级数*.

不应该认为每个无穷可微的函数的泰勒级数都在点 x_0 的某个邻域内收敛, 因为对于任何一个数列 $c_0, c_1, \cdots, c_n, \cdots$ 都可以构造 (这并不是很简单的) 一个函数 $f(x)$, 使得 $f^{(n)}(x_0) = c_n, n \in \mathbb{N}$.

也不应该认为, 如果泰勒级数收敛, 它就一定收敛到产生它的函数. 泰勒级数收敛到产生它的函数, 此事只对于所谓*解析函数*成立.

下面就是一个柯西给出的非解析函数的例子:

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & \text{若} x \neq 0, \\ 0, & \text{若} x = 0. \end{cases}$$

从导数的定义出发, 由于当 $x \to 0$ 时, 不管 k 值如何总有 $x^k e^{-\frac{1}{x^2}} \to 0$ (见第三章 §2 的例 30), 可以验证对于 $n = 0, 1, 2, \cdots, f^{(n)}(0) = 0$. 于是, 在这种情况下泰勒级数的每一项都是零, 从而它的和恒等于零, 同时, 当 $x \neq 0$ 时, $f(x) \neq 0$.

最后我们来详细研究一下泰勒公式的局部形式.

重新回到用多项式局部逼近函数 $f: E \to \mathbb{R}$ 的问题. 我们是在 §1 第 3 段中开始讨论这个问题的. 我们试图选择一个多项式 $P_n(x_0; x) = c_0 + c_1(x-x_0) + \cdots + c_n(x-x_0)^n$, 使得

$$f(x) = P_n(x_0; x) + o((x-x_0)^n), \text{当} x \to x_0, x \in E,$$

或详细写出来

$$f(x)=c_0+c_1(x-x_0)+\cdots+c_n(x-x_0)^n+o((x-x_0)^n),$$
$$\text{当} x\to x_0, x\in E. \tag{32}$$

我们明确地叙述一个实质上已经证明了的

命题 3 *如果满足条件 (32) 的多项式*

$$P_n(x_0;x)=c_0+c_1(x-x_0)+\cdots+c_n(x-x_0)^n$$

存在, 那么它是唯一的.

◄ 实际上多项式的系数可以从条件 (32) 顺次完全单值地 (根据极限的唯一性) 求出来:

$$\begin{aligned}
c_0 &= \lim_{E\ni x\to x_0} f(x),\\
c_1 &= \lim_{E\ni x\to x_0} \frac{f(x)-c_0}{x-x_0},\\
&\cdots\cdots\cdots\cdots\\
c_n &= \lim_{E\ni x\to x_0} \frac{f(x)-[c_0+\cdots+c_{n-1}(x-x_0)^{n-1}]}{(x-x_0)^n}.
\end{aligned}$$ ►

现在证明

命题 4 (局部泰勒公式) *设 E 是以 $x_0\in\mathbb{R}$ 为端点的闭区间. 如果函数 $f:E\to\mathbb{R}$ 在点 x_0 处有直到 n 阶的导数 $f'(x_0),\cdots,f^{(n)}(x_0)$, 那么下列表达式成立:*

$$\begin{aligned}
f(x)= {} & f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n\\
&+o((x-x_0)^n), \text{当} x\to x_0, x\in E.
\end{aligned}\tag{33}$$

这样一来, 可微函数的局部逼近问题就由相应阶数的泰勒多项式解决了.

由于泰勒多项式$P_n(x_0;x)$ 是根据这样的条件构造的, 即在点 x_0 处它和函数 f 的 n 阶以内的同阶导数相同, 所以 $f^{(k)}(x_0)-P_n^{(k)}(x_0;x_0)=0(k=0,1,\cdots,n)$, 从而公式 (33) 的正确性由下述引理确立.

引理 2 *若函数 $\varphi:E\to\mathbb{R}$ 定义在以 x_0 为端点的闭区间上, 在点 x_0 处有一切不超过 n 阶的导数 $\varphi'(x_0),\cdots,\varphi^{(n)}(x_0)$, 且 $\varphi(x_0)=\varphi'(x_0)=\cdots=\varphi^{(n)}(x_0)=0$, 则当 $x\to x_0, x\in E$时, $\varphi(x)=o((x-x_0)^n)$.*

◄ 当 $n=1$ 时, 结论可从函数 φ 在点 x_0 处可微的定义如下推出: 据可微的定义

$$\varphi(x)=\varphi(x_0)+\varphi'(x_0)(x-x_0)+o(x-x_0), \text{当} x\to x_0, x\in E.$$

但因 $\varphi(x_0)=\varphi'(x_0)=0$, 故有

$$\varphi(x)=o(x-x_0), \text{当} x\to x_0, x\in E \text{时}.$$

现在假定命题对阶数 $n=k-1\geqslant 1$ 已证实. 我们来证它对于阶数 $n=k\geqslant 2$ 也成立.

我们预先指出, 由于

$$\varphi^{(k)}(x_0)=(\varphi^{(k-1)})'(x_0)=\lim_{E\ni x\to x_0}\frac{\varphi^{(k-1)}(x)-\varphi^{(k-1)}(x_0)}{x-x_0},$$

那么, $\varphi^{(k)}(x_0)$ 的存在是以 $\varphi^{(k-1)}(x)$ 在 E 上 x_0 的一个不管多么小的邻域中有定义为前提的. 如果必要的话, 只消缩短闭区间 E, 就可以预先认为函数 $\varphi(x)$、$\varphi'(x), \cdots$, $\varphi^{(k-1)}(x)$, 其中 $k\geqslant 2$, 在以 x_0 为端点的整个闭区间 E 上有定义. 由于 $k\geqslant 2$, 那么, 所以函数 $\varphi(x)$ 在 E 上有导数 $\varphi'(x)$, 且根据条件, 有

$$(\varphi')'(x_0)=\cdots=(\varphi')^{(k-1)}(x_0)=0.$$

这样一来, 根据归纳假设

$$\varphi'(x)=o((x-x_0)^{k-1}) \text{ 当} x\to x_0, x\in E \text{时}.$$

此时, 用拉格朗日定理, 得到

$$\begin{aligned}\varphi(x)&=\varphi(x)-\varphi(x_0)=\varphi'(\xi)(x-x_0)\\&=\alpha(\xi)(\xi-x_0)^{k-1}(x-x_0),\end{aligned}$$

其中 ξ 是位于 x_0 和 x 之间的点, 即 $|\xi-x_0|<|x-x_0|$, 而当 $\xi\to x_0$、$\xi\in E$ 时, $\alpha(\xi)\to 0$. 这就意味着, 当 $x\to x_0$、$x\in E$ 时, 同时有 $\xi\to x_0$、$\xi\in E$ 且 $\alpha(\xi)\to 0$. 而由于

$$|\varphi(x)|\leqslant|\alpha(\xi)||x-x_0|^{k-1}|x-x_0|,$$

所以, 就证明了

$$\varphi(x)=o((x-x_0)^k) \quad x\to x_0, x\in E.$$

这样一来, 根据数学归纳法的原理, 引理 2 的结论得到了证明. ►

关系式 (33) 之所以叫做*局部泰勒公式*, 是因为根据其中余项的形式 (所谓*佩亚诺形式*)

$$r_n(x_0;x)=o((x-x_0)^n) \tag{34}$$

只能对泰勒多项式和当 $x\to x_0; x\in E$ 时函数的渐近行为作出结论.

因此, 公式 (33) 对于计算极限及描述当 $x\to x_0$、$x\in E$ 时函数的渐近行为是适用的, 但是, 只要对量 $r_n(x_0;x)=o((x-x_0)^n)$ 还没有作出实际的估计, 它都不能应用于函数值的近似计算.

总结一下. 我们定义了泰勒多项式

$$P_n(x_0;x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n,$$

写出了泰勒公式

$$f(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+r_n(x_0;x)$$

并得到了它的下列重要的具体表达式:

如果 f 在以 x_0、x 为端点的开区间中有 $n+1$ 阶导数, 则

$$f(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}, \quad (35)$$

其中 ξ 是位于 x_0 和 x 之间的点;

如果 f 在点 x_0 处有直到 $n\geqslant 1$ 阶的导数, 则

$$f(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+o((x-x_0)^n), \quad (36)$$

关系式 (35) 称为*具有拉格朗日余项的泰勒公式*, 它显然是拉格朗日定理的推广, 后者是它当 $n=0$ 时的特例.

关系式 (36) 称为*具有佩亚诺余项的泰勒公式*, 它显然是函数在一点处可微的定义的推广, 后者是它当 $n=1$ 时的特例.

我们注意到, 公式 (35) 实际上含有更多的内容, 因为一方面如我们已见到的, 用它能够估计余项的绝对值, 而另一方面, 例如, 当 $f^{(n+1)}(x)$ 在 x_0 的邻域内有界时, 从它还推出渐近公式

$$f(x)=f(x_0)+\frac{f'(x_0)}{1!}(x-x_0)+\cdots+\frac{f^{(n)}(x_0)}{n!}(x-x_0)^n+O((x-x_0)^{n+1}). \quad (37)$$

因此, 对于无穷可微的函数, 公式 (35) 包含着局部公式 (36), 而经典分析在绝大多数场合都是与这种函数打交道的.

特别地, 根据公式 (37) 和前面研究过的例3—例10, 现在可以写出下面的当

$x \to 0$ 时的渐近公式表:

$$e^x = 1 + \frac{1}{1!}x + \cdots + \frac{1}{n!}x^n + O(x^{n+1}),$$

$$\cos x = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \cdots + \frac{(-1)^n}{(2n)!}x^{2n} + O(x^{2n+2}),$$

$$\sin x = x - \frac{1}{3!}x^3 + \cdots + \frac{(-1)^n}{(2n+1)!}x^{2n+1} + O(x^{2n+3}),$$

$$\operatorname{ch} x = 1 + \frac{1}{2!}x^2 + \frac{1}{4!}x^4 + \cdots + \frac{1}{(2n)!}x^{2n} + O(x^{2n+2}),$$

$$\operatorname{sh} x = x + \frac{1}{3!}x^3 + \cdots + \frac{1}{(2n+1)!}x^{2n+1} + O(x^{2n+3}),$$

$$\ln(1+x) = x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \cdots + \frac{(-1)^{n-1}}{n}x^n + O(x^{n+1}),$$

$$(1+x)^\alpha = 1 + \frac{\alpha}{1!}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \cdots + \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}x^n + O(x^{n+1}).$$

现在再考察几个应用泰勒公式的例子.

例 11 写出计算函数 $\sin x$ 在闭区间 $-1 \leqslant x \leqslant 1$ 上的值的多项式, 使绝对误差不超过 10^{-3}.

我们取函数 $\sin x$ 在点 $x_0 = 0$ 的邻域内展开所得适当次数的泰勒多项式作为这样的多项式. 因为

$$\sin x = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \cdots + \frac{(-1)^n}{(2n+1)!}x^{2n+1} + 0 \cdot x^{2n+2} + r_{2n+2}(0;x),$$

其中余项按拉格朗日公式给出:

$$r_{2n+2}(0;x) = \frac{\sin\left(\xi + \frac{\pi}{2}(2n+3)\right)}{(2n+3)!}x^{2n+3}.$$

所以当 $|x| \leqslant 1$ 时,

$$|r_{2n+2}(0;x)| \leqslant \frac{1}{(2n+3)!}.$$

但当 $n \geqslant 2$ 时, $\frac{1}{(2n+3)!} < 10^{-3}$, 所以

$$\sin x \approx x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5,$$

它在闭区间 $|x| \leqslant 1$ 上具有所需的精确性.

例 12 我们证明当 $x \to 0$ 时, $\operatorname{tg} x = x + \frac{1}{3}x^3 + o(x^3)$.

我们有

$$\operatorname{tg}' x = \cos^{-2} x, \quad \operatorname{tg}'' x = 2\cos^{-3} x \cdot \sin x,$$
$$\operatorname{tg}''' x = 6\cos^{-4} x \sin^2 x + 2\cos^{-2} x.$$

于是, $\operatorname{tg} 0 = 0, \operatorname{tg}' 0 = 1, \operatorname{tg}'' 0 = 0, \operatorname{tg}''' 0 = 2$, 从而由局部泰勒公式推出所需的关系式.

例 13　设 $\alpha > 0$. 我们来研究级数 $\sum_{n=1}^{\infty} \ln \cos \frac{1}{n^\alpha}$ 的收敛性.

当 $\alpha > 0, n \to \infty$ 时, $\frac{1}{n^\alpha} \to 0$. 估计级数项的阶

$$\ln \cos \frac{1}{n^\alpha} = \ln \left(1 - \frac{1}{2!} \frac{1}{n^{2\alpha}} + o\left(\frac{1}{n^{2\alpha}}\right)\right) = -\frac{1}{2} \frac{1}{n^{2\alpha}} + o\left(\frac{1}{n^{2\alpha}}\right).$$

于是我们的级数是常号级数, 其项与级数 $\sum_{n=1}^{\infty} -\frac{1}{2} \frac{1}{n^{2\alpha}}$ 的项等价. 由于后者仅当 $\alpha > \frac{1}{2}$ 时收敛, 所以在所说的区域 $\alpha > 0$ 中, 原来的级数仅当 $\alpha > \frac{1}{2}$ 时收敛 (见问题 15b)).

例 14　我们证明当 $x \to 0$ 时,

$$\ln \cos x = -\frac{1}{2}x^2 - \frac{1}{12}x^4 - \frac{1}{45}x^6 + o(x^6).$$

这次我们不去逐次计算六个导数, 而是使用已经知道的当 $x \to 0$ 时, $\cos x$ 的展开式和当 $u \to 0$ 时, $\ln(1+u)$ 的展开式:

$$\begin{aligned}
\ln \cos x &= \ln \left(1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + O(x^8)\right) \\
&= \ln(1+u) = u - \frac{1}{2}u^2 + \frac{1}{3}u^3 + O(u^4) \\
&= \left(-\frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + O(x^8)\right) \\
&\quad -\frac{1}{2}\left(\frac{1}{(2!)^2}x^4 - 2 \times \frac{1}{2!}\frac{1}{4!}x^6 + O(x^8)\right) \\
&\quad +\frac{1}{3}\left(-\frac{1}{2^3}x^6 + O(x^8)\right) \\
&= -\frac{1}{2}x^2 - \frac{1}{12}x^4 - \frac{1}{45}x^6 + O(x^8).
\end{aligned}$$

例 15　求函数 $\ln \cos x$ 在 $x = 0$ 处的前六个导数的值.

$(\ln \cos x)' = \frac{-\sin x}{\cos x}$. 于是在 $x = 0$ 处, 由于 $\cos 0 \neq 0$, 所以这个函数显然有任意阶的导数. 我们不从求这些导数的函数表达式着手, 而是使用泰勒多项式的唯一性和上例的结果.

若

$$f(x)=c_0+c_1x+\cdots+c_nx^n+o(x^n)\quad 当x\to 0,$$

则

$$c_k=\frac{f^{(k)}(0)}{k!}, 即f^{(k)}(0)=k!c_k.$$

于是, 在我们的情形, 有

$$(\ln\cos)(0)=0,\quad (\ln\cos)'(0)=0,\quad (\ln\ \cos)''(0)=-\frac{1}{2}\cdot 2!,$$
$$(\ln\cos)^{(3)}(0)=0,\quad (\ln\ \cos)^{(4)}(0)=-\frac{1}{12}\cdot 4!,$$
$$(\ln\cos)^{(5)}(0)=0,\quad (\ln\ \cos)^{(6)}(0)=-\frac{1}{45}\cdot 6!.$$

例 16 设 $f(x)$ 是在点 $x_0=0$ 处无穷可微的函数, 并已知它的导数在零的邻域内的展开式

$$f'(x)=c_0'+c_1'x+\cdots+c_n'x^n+O(x^{n+1}).$$

那么, 根据泰勒展开式的唯一性, 我们有

$$(f')^{(k)}(0)=k!c_k',$$

因此 $f^{(k+1)}(0)=k!c_k'$. 这样一来, 对于函数 $f(x)$ 本身, 有展开式

$$f(x)=f(0)+\frac{c_0'}{1!}x+\frac{1!c_1'}{2!}x^2+\cdots+\frac{n!c_n'}{(n+1)!}x^{n+1}+O(x^{n+2}),$$

化简之后即

$$f(x)=f(0)+\frac{c_0'}{1}x+\frac{c_1'}{2}x^2+\cdots+\frac{c_n'}{n+1}x^{n+1}+O(x^{n+2}).$$

例 17 求函数 $f(x)=\operatorname{arctg}x$ 在零处的泰勒展开式.

由于 $f'(x)=\dfrac{1}{1+x^2}=(1+x^2)^{-1}=1-x^2+x^4-\cdots+(-1)^nx^{2n}+O(x^{2n+2})$, 所以根据上例所说的理由,

$$f(x)=f(0)+\frac{1}{1}x-\frac{1}{3}x^3+\frac{1}{5}x^5-\cdots$$
$$+\frac{(-1)^n}{2n+1}x^{2n+1}+O(x^{2n+3}),$$

即

$$\operatorname{arctg}x=x-\frac{1}{3}x^3+\frac{1}{5}x^5-\cdots+\frac{(-1)^n}{2n+1}x^{2n+1}+O(x^{2n+3}).$$

例 18 类似地, 把函数 $\arcsin' x = (1-x^2)^{-\frac{1}{2}}$ 在零的邻域内按泰勒公式展开, 逐次求出

$$(1+u)^{-\frac{1}{2}} = 1 + \frac{-\frac{1}{2}}{1!}u + \frac{-\frac{1}{2}\left(-\frac{1}{2}-1\right)}{2!}u^2 + \cdots + \frac{-\frac{1}{2}\left(-\frac{1}{2}-1\right)\cdot\cdots\cdot\left(-\frac{1}{2}-n+1\right)}{n!}u^n + O(u^{n+1}),$$

$$(1-x^2)^{-\frac{1}{2}} = 1 + \frac{1}{2}x^2 + \frac{1\cdot 3}{2^2\cdot 2!}x^4 + \cdots + \frac{1\cdot 3\cdot\cdots\cdot(2n-1)}{2^n\cdot n!}x^{2n} + O(x^{2n+2}),$$

$$\arcsin x = x + \frac{1}{2\cdot 3}x^3 + \frac{1\cdot 3}{2^2\cdot 2!\cdot 5}x^5 + \cdots + \frac{(2n-1)!!}{(2n)!!(2n+1)}x^{2n+1} + O(x^{2n+3}),$$

或在经过一些初等变换后得

$$\arcsin x = x + \frac{1}{3!}x^3 + \frac{(3!!)^2}{5!}x^5 + \cdots + \frac{[(2n-1)!!]^2}{(2n+1)!}x^{2n+1} + O(x^{2n+3}).$$

这里, $(2n-1)!! = 1\cdot 3\cdot\cdots\cdot(2n-1)$, $(2n)!! = 2\cdot 4\cdot\cdots\cdot(2n)$.

例 19 利用例 5、例 12、例 17、例 18 的结果求得

$$\begin{aligned}
&\lim_{x\to 0}\frac{\operatorname{arctg} x - \sin x}{\operatorname{tg} x - \arcsin x} \\
&= \lim_{x\to 0}\frac{\left[x - \frac{1}{3}x^3 + O(x^5)\right] - \left[x - \frac{1}{3!}x^3 + O(x^5)\right]}{\left[x + \frac{1}{3}x^3 + O(x^5)\right] - \left[x + \frac{1}{3!}x^3 + O(x^5)\right]} \\
&= \lim_{x\to 0}\frac{-\frac{1}{6}x^3 + O(x^5)}{\frac{1}{6}x^3 + O(x^5)} = -1.
\end{aligned}$$

练　习

1. 选择数 a 和 b 使得函数 $f(x) = \cos x - \dfrac{1+ax^2}{1+bx^2}$ 当 $x\to 0$ 时是尽可能高阶的无穷小量.
2. 求 $\lim\limits_{x\to\infty} x\left[\dfrac{1}{e} - \left(\dfrac{x}{x+1}\right)^x\right]$.
3. 写出函数 e^x 在零点处的一个泰勒多项式, 使能按它在闭区间 $-1\leqslant x\leqslant 2$ 上计算 e^x 的值精确到 10^{-3}.
4. 设 f 是在零点无穷可微的函数. 证明:
   a) 若 f 是偶函数, 则它在零这点的泰勒级数只含有 x 的偶次幂.
   b) 若 f 是奇函数, 则它在零这点的泰勒级数只含有 x 的奇次幂.

5. 证明: 若 $f \in C^{(\infty)}_{[-1,1]}$, $f^{(n)}(0)=0$, 当 $n=0,1,2,\cdots$, 并且存在数 C, 使得

$$\sup_{-1\leqslant x\leqslant 1}|f^{(n)}(x)| \leqslant n!C^n, \quad n\in\mathbf{N},$$

那么, 在 $[-1,1]$ 上 $f(x)\equiv 0$.

6. 设 $f\in C^{(n)}(]-1,1[)$ 且 $\sup\limits_{-1<x<1}|f(x)|\leqslant 1$.

设 $m_k(I)=\inf\limits_{x\in I}|f^{(k)}(x)|$, 其中 I 是含在开区间 $]-1,1[$ 中的区间. 证明:

a) 若 I 被分成顺次的三个区间 I_1、I_2、I_3, 且 μ 是 I_2 的长, 则

$$m_k(I)\leqslant \frac{1}{\mu}(m_{k-1}(I_1)+m_{k-1}(I_3)).$$

b) 若 I 的长是 λ, 则

$$m_k(I)\leqslant \frac{2^{\frac{k(k+1)}{2}}k^k}{\lambda^k}.$$

c) 存在这样的数 α_n, 它只与 n 有关, 使得若 $|f'(0)|\geqslant\alpha_n$, 则方程

$$f^{(n)}(x)=0$$

在 $]-1,1[$ 中至少有 $n-1$ 个不同的根.

提示 在 b) 中用 a) 和归纳法原理; 在 c) 中用 a) 并归纳地证明: 在 $]-1,1[$ 中存在这样的点列

$$x_{k_1}<x_{k_2}<\cdots<x_{k_k},$$

使得对于 $1\leqslant i\leqslant k-1$, $f^{(k)}(x_{k_i})\cdot f^{(k)}(x_{k_{i+1}})<0$.

7. 证明: 若函数 f 在开区间 I 上定义且可微, $[a,b]\subset I$, 则

a) 函数 $f'(x)$(不必是连续的!) 在 $[a,b]$ 上取遍 $f'(a)$ 和 $f'(b)$ 之间的一切值 (这是达布[①]定理).

b) 若还有 $f''(x)$ 在 $]a,b[$ 中存在, 则存在点 $\xi\in]a,b[$, 使得

$$f'(b)-f'(a)=f''(\xi)(b-a).$$

8. 函数 $f(x)$ 在全数轴上可微且 $f'(x)$ 可以不连续 (见 §1 第 5 段例 7).

a) 证明: 函数 $f'(x)$ 只可能有第二类间断点.

b) 指出下述对于 $f'(x)$ 的连续性的 "证明" 中的错误.

◀ 设 x_0 是 $\mathbb{R}$ 中任意一点且 $f'(x_0)$ 是函数 f 在点 x_0 处的导数. 根据导数的定义和拉格朗日定理

$$f'(x_0)=\lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}=\lim_{x\to x_0}f'(\xi)=\lim_{\xi\to x_0}f'(\xi),$$

其中 ξ 是 x_0 和 x 之间的点, 从而当 $x\to x_0$ 时也趋于 x_0. ▶

---

①达布 (G.Darboux)(1842—1917)—— 法国数学家.

9. 设 f 是区间 I 上的二次可微函数. 设

$$M_0 = \sup_{x\in I}|f(x)| \quad M_1 = \sup_{x\in I}|f'(x)|, M_2 = \sup_{x\in I}|f''(x)|.$$

证明:

a) 若 $I = [-a, a]$, 则

$$|f'(x)| \leqslant \frac{M_0}{a} + \frac{x^2+a^2}{2a}M_2.$$

b)

$$\begin{cases} M_1 \leqslant 2\sqrt{M_0M_2}, \text{若}I\text{的长度不小于}2\sqrt{\dfrac{M_0}{M_1}}, \\ M_1 \leqslant \sqrt{2M_0M_2}, \text{若}I = \mathbb{R}. \end{cases}$$

c) 在 b) 中数 2 和 $\sqrt{2}$ 不可再减小.

d) 若 f 在 $\mathbb{R}$ 中 p 次可微, 且量 M_0 和 $M_p = \sup\limits_{x\in\mathbb{R}}|f^{(p)}(x)|$ 都有限, 则对于 $1 \leqslant k \leqslant p$, 量 $M_k = \sup\limits_{x\in\mathbb{R}}|f^{(k)}(x)|$ 亦有限且

$$M_k \leqslant 2^{\frac{k(p-k)}{2}}M_0^{1-\frac{k}{p}}M_p^{\frac{k}{p}}.$$

提示 利用问题 6b),9b) 和归纳法原理.

10. 证明: 若函数 f 在点 x_0 处有直到 $n+1$ 阶导数, 则泰勒公式的拉格朗日型余项

$$r_n(x_0;x) = \frac{1}{n!}f^{(n)}(x_0+\theta(x-x_0))(x-x_0)^n,$$

其中的量 $\theta = \theta(x)$ 当 $x \to x_0$ 时趋于 $\dfrac{1}{n+1}$.

11. 设 f 是区间 I 上的 n 次可微函数. 证明:

a) 若 f 在区间 I 的 $n+1$ 个点处为零, 则存在点 $\xi \in I$ 使 $f^{(n)}(\xi) = 0$.

b) 如果 $x_1, x_2, \cdots, x_p$ 是区间 I 的点, 则存在唯一的不超过 $n-1$ 阶的多项式 $L(x)$(拉格朗日插值多项式), 使得 $f(x_i) = L(x_i)(i = 1, 2, \cdots, n)$. 此外, 对于 $x \in I$ 存在 $\xi \in I$ 使得

$$f(x) - L(x) = \frac{(x-x_1)\cdot\cdots\cdot(x-x_n)}{n!}f^{(n)}(\xi).$$

c) 若 $x_1 < x_2 < \cdots < x_p$ 是区间 I 的点, $n_i(1 \leqslant i \leqslant p)$ 是这样一些自然数, 使得 $n_1 + n_2 + \cdots + n_p = n$ 且当 $0 \leqslant k \leqslant n_i - 1$ 时, $f^{(k)}(x_i) = 0$, 则在区间 $[x_1, x_p]$ 中存在点 ξ 使 $f^{(n-1)}(\xi) = 0$.

d) 存在唯一的 $n-1$ 阶多项式 $H(x)$(埃尔米特[①]插值多项式), 使得当 $0 \leqslant k \leqslant n_i - 1$ 时 $f^{(n)}(x_i) = P^{(k)}(x_i)$. 此外, 在包含点 x 和 $x_i(i = 1, \cdots, p)$ 的最小的区间内部存在点 ξ 使得

$$f(x) = H(x) + \frac{(x-x_1)^{n_1}\cdot\cdots\cdot(x-x_p)^{n_p}}{n!}f^{(n)}(\xi).$$

①埃尔米特 (Hermite)(1822—1901) 法国数学家, 从事分析学问题的研究, 特别是证明了数 e 的超越性.

这个公式叫做埃尔米特插值公式. 诸点 $x_i(i=1,\cdots,p)$ 分别叫做n_i 重插值节点. 拉格朗日插值公式 (问题 b) 是埃尔米特公式当 $p=n, n_i=1(i=1,\cdots,n)$ 时的特例, 同样, 具有拉格朗日余项的泰勒公式也是埃尔米特插值公式的特例, 它相应于 $p=1$ 的情形, 即在一个 n 重插值点上插值的情形.

12. 证明:

a) 在实系数多项式 $P(x)$ 的两个实根之间存在它的导数 $P'(x)$ 的根.

b) 若多项式 $P(x)$ 有重根, 则多项式 $P'(x)$ 也有这个根, 但重数少 1.

c) 若 $Q(x)$ 是多项式 $P(x)$ 和 $P'(x)$ 的最大公因式, 其中 $P'(x)$ 是多项式 $P(x)$ 的导数, 则多项式 $\dfrac{P(x)}{Q(x)}$ 以多项式 $P(x)$ 的根为根, 且都是一重的.

13. 证明:

a) 任一多项式 $P(x)$ 都可表成 $c_0+c_1(x-x_0)+\cdots+c_n(x-x_0)^n$ 的形式.

b) 存在唯一的 n 阶多项式 $P(x)$ 使得当 $E\ni x\to x_0$ 时,

$$f(x)-P(x)=o((x-x_0)^n).$$

这里 f 是定义在集 E 上的函数, 而 x_0 是 E 的极限点.

14. 借助于对 $k, 1\leqslant k$ 的归纳法, 我们定义函数 f 在点 x_0 处的 k 阶有限差:

$$\begin{aligned}\Delta^1 f(x_0;h_1)&=\Delta f(x_0;h_1)=f(x_0+h_1)-f(x_0),\\ \Delta^2 f(x_0;h_1,h_2)&=\Delta\Delta f(x_0;h_1,h_2)\\ &=(f(x_0+h_1+h_2)-f(x_0+h_2))-(f(x_0+h_1)-f(x_0))\\ &=f(x_0+h_1+h_2)-f(x_0+h_1)-f(x_0+h_2)+f(x_0),\\ &\cdots\cdots\cdots\cdots\\ \Delta^k f(x_0;h_1,\cdots,h_k)&=\Delta^{k-1}g_k(x_0;h_1,\cdots,h_{k-1}),\end{aligned}$$

其中 $g_k(x)=\Delta^1 f(x;h_k)=f(x+h_k)-f(x)$.

a) 设 $f\in C^{(n-1)}[a,b]$, 且至少在开区间 $]a,b[$ 中 $f^{(n)}(x)$ 存在. 若诸点

$$x_0, x_0+h_1, x_0+h_2, x_0+h_1+h_2,\cdots,x_0+h_1+\cdots+h_n$$

皆在 $[a,b]$ 中, 则在包含这些点的最小闭区间的内部存在点 ξ 使得

$$\Delta^n f(x_0;h_1,\cdots,h_n)=f^{(n)}(\xi)h_1\cdot\cdots\cdot h_n.$$

b) (续) 若 $f^{(n)}(x_0)$ 存在, 则成立估计式

$$\begin{aligned}&|\Delta^n f(x_0;h_1,\cdots,h_n)-f^{(n)}(x_0)h_1\cdot\cdots\cdot h_n|\\ &\leqslant\sup_{x\in]a,b[}|f^{(n)}(x)-f^{(n)}(x_0)|\cdot|h_1|\cdot\cdots\cdot|h_n|.\end{aligned}$$

c) (续) 置 $\Delta^n f(x_0;h,\cdots,h)=\Delta^n f(x_0;h^n)$. 证明, 若 $f^{(n)}(x_0)$ 存在, 则

$$f^{(n)}(x_0)=\lim_{h\to 0}\frac{\Delta^n f(x_0;h^n)}{h^n}.$$

d) 举例证明, 即使 $f^{(n)}(x_0)$ 不存在, 上面的极限也可以存在.

提示 例如对于函数

$$f(x)=\begin{cases}x^3\sin\frac{1}{x}, & x\neq 0,\\ 0, & x=0,\end{cases}$$

考察 $\Delta^2 f(0;h^2)$ 并证明

$$\lim_{h\to 0}\frac{\Delta^2 f(0;h^2)}{h^2}=0.$$

15. a) 将拉格朗日定理用于函数 $\frac{1}{x^\alpha}$, 其中 $\alpha>0$, 证明对于 $n\in\mathbb{N}$ 和 $\alpha>0$ 成立不等式

$$\frac{1}{n^{1+\alpha}}<\frac{1}{\alpha}\left(\frac{1}{(n-1)^\alpha}-\frac{1}{n^\alpha}\right).$$

b) 用问题 a) 的结果证明级数 $\sum\limits_{n=1}^{\infty}\frac{1}{n^\sigma}$ 当 $\sigma>1$ 时收敛.

§4. 用微分学的方法研究函数

1. 函数单调的条件(参看函数单调性检验法)

命题 1 *开区间 $]a,b[=E$ 上可微的函数 $f:E\to\mathbb{R}$ 在此区间内的单调性与它的导数 f' 在此区间的符号 (正、负) 彼此之间有下述的关系:*

$f'(x)>0\Rightarrow f$ 递增 $\Rightarrow f'(x)\geqslant 0$,
$f'(x)\geqslant 0\Rightarrow f$ 不减 $\Rightarrow f'(x)\geqslant 0$,
$f'(x)\equiv 0\Rightarrow f\equiv \mathrm{const}$(常数) $\Rightarrow f'(x)\equiv 0$,
$f'(x)\leqslant 0\Rightarrow f$ 不增 $\Rightarrow f'(x)\leqslant 0$,
$f'(x)<0\Rightarrow f$ 递减 $\Rightarrow f'(x)\leqslant 0$.

◀ 左边这栏的结论我们在讨论拉格朗日定理时已经知道了, 根据拉格朗日定理

$$f(x_2)-f(x_1)=f'(\xi)(x_2-x_1),$$

其中 $x_1,x_2\in]a,b[$ 且 ξ 在 x_1 和 x_2 之间. 从这个公式见到, 当 $x_1<x_2$ 时, 差 $f(x_2)-f(x_1)$ 的正负与 $f'(\xi)$ 的正负一致.

右边那栏的结论直接由导数的定义得出. 例如我们来证明, 若在 $]a,b[$ 上可微的函数递增, 则在 $]a,b[$ 上 $f'(x)\geqslant 0$. 实际上

$$f'(x)=\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}.$$

若 $h>0$, 则 $f(x+h)-f(x)>0$, 而若 $h<0$, 则 $f(x+h)-f(x)<0$; 因此极限号后面的分数是正的.

因此, 它的极限 $f'(x)$ 非负, 这就是所要的结论. ▶

注 1 从函数 $f(x) = x^3$ 这个例子见到, 可微函数的递增只蕴含导数的非负而不蕴含导数恒正. 在这个例中, $f'(0) = 3x^2|_{x=0} = 0$.

注 2 在 $A \Rightarrow B$ 这个记号中, 如我们已经及时指出过的, A 是 B 的充分条件而 B 是 A 的必要条件. 这就意味着, 从命题 1 特别地可以做出下面的结论:

函数在开区间上为常数当且仅当它的导数在这个区间上恒等于零;

要使在开区间上可微的函数在这个区间上递减, 只需使它的导数在这个区间的任何一点处都是负的;

为使在开区间上可微的函数在这个区间上递减, 它的导数在这个区间上必须是非正的.

例 1 设在 $\mathbb{R}$ 上 $f(x) = x^3 - 3x + 2$. 那么 $f'(x) = 3x^2 - 3 = 3(x^2 - 1)$, 并且, 由于当 $|x| < 1$ 时, $f'(x) < 0$, 以及当 $|x| > 1$ 时, $f'(x) > 0$, 所以可以断言, 函数在开区间 $]-\infty, -1[$ 上递增, 在开区间 $]-1, 1[$ 上递减, 而在开区间 $]1, +\infty[$ 上重又递增.

2. 函数的内极值点条件 根据费马引理 (§3. 引理 1), 可以得到下面的

命题 2 (内极值点的必要条件) *要使点 x_0 是定义在这点的邻域中的函数 $f : U(x_0) \to \mathbb{R}$ 的极值点, 必须成立下列两条件之一: 或者函数在 x_0 不可微, 或者 $f'(x_0) = 0$.*

简单的例子表明, 极值点的这些必要条件不是充分的.

例 2 设在 $\mathbb{R}$ 上 $f(x) = x^3$. 那么 $f'(0) = 0$, 但在点 $x_0 = 0$ 处没有极值.

例 3 设

$$f(x) = \begin{cases} x, & \text{当} x \geqslant 0, \\ 2x, & \text{当} x < 0. \end{cases}$$

这个在零点发生转折的函数显然在这点既无导数亦无极值.

例 4 我们来求函数 $f(x) = x^2$ 在闭区间 $[-2, 1]$ 上的最大值. 在这种情况下, 显然最大值在区间 -2 的一端处达到. 但正规的求法是这样的. 求出 $f'(x) = 2x$ 以及开区间 $]-2, 1[$ 中一切使 $f'(x) = 0$ 的点. 在我们的题目中, 这样的点, 只有一个, 就是点 $x = 0$. $f(x)$ 的最大值或在这些点中达到, 或在一个端点处达到. 关于端点的情形, 命题 2 什么也没说. 因此, 在我们的情况下, 应当把函数值 $f(-2) = 4, f(0) = 0, f(1) = 1$ 进行比较, 由此断定函数 $f(x) = x^2$ 在闭区间上的最大值等于 4, 它是在这个区间的一端 -2 处达到的.

根据在第 1 段中建立的关于导数的符号和函数的单调性之间的联系, 我们得到下述在一点处存在局部极值或不存在局部极值的充分条件.

命题 3 (用一阶导数表达的极值的充分条件) 设$f: U(x_0) \to \mathbb{R}$ 是定义在点 x_0 的邻域内的函数, 在点 x_0 处连续且在它的空心邻域 $\overset{\circ}{U}(x_0)$ 中可微. 设

$$\overset{\circ}{U}^-(x_0) = \{x \in U(x_0) | x < x_0\}, \quad \overset{\circ}{U}^+(x_0) = \{x \in U(x_0) | x_0 < x\}.$$

那么下列断语成立:

a) $(\forall x \in \overset{\circ}{U}^-(x_0)(f'(x) < 0)) \wedge (\forall x \in \overset{\circ}{U}^+(x_0)(f'(x) < 0))$
$\Rightarrow$ (f在 x_0处没有极值);

b) $(\forall x \in \overset{\circ}{U}^-(x_0)(f'(x) < 0)) \wedge (\forall x \in \overset{\circ}{U}^+(x_0)(f'(x) > 0))$
$\Rightarrow$ (x_0是 f的严格局部极小值点);

c) $(\forall x \in \overset{\circ}{U}^-(x_0)(f'(x) > 0)) \wedge (\forall x \in \overset{\circ}{U}^+(x_0)(f'(x) < 0))$
$\Rightarrow$ (x_0是 f的严格局部极大值点);

d) $(\forall x \in \overset{\circ}{U}^-(x_0)(f'(x) > 0)) \wedge (\forall x \in \overset{\circ}{U}^+(x_0)(f'(x) > 0))$
$\Rightarrow$ (f在 x_0处没有极值).

可以简单地但不太确切地说, 如果经过一个点时导数变号, 则有极值, 而若导数不变号, 则没有极值.

我们立即注意到, 关于极值的这些条件是充分的, 但不是必要的. 下面的例子可以说明这一点.

例 5 设

$$f(x) = \begin{cases} 2x^2 + x^2 \sin \dfrac{1}{x}, & 当x \neq 0, \\ 0, & 当x = 0. \end{cases}$$

由于 $x^2 \leqslant f(x) \leqslant 3x^2$, 那么很明显, 函数在点 $x_0 = 0$ 处有严格局部极小值. 但无论在这个点的哪个空心半邻域中, 它的导数 $f'(x) = 4x + 2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x}$ 都不保号. 这样一个例子还指出了由于命题 3 的上述简略表达而可能引起的误会.

现在证明命题 3.

◀ a) 从命题 2 推出函数 f 在 $\overset{\circ}{U}^-(x_0)$ 上严格递减. 由于 f 在 x_0 连续, 故

$$\lim_{\overset{\circ}{U}^-(x_0) \ni x \to x_0} f(x) = f(x_0),$$

从而当 $x \in \overset{\circ}{U}^-(x_0)$ 时, $f(x) > f(x_0)$. 同理, 当 $x \in \overset{\circ}{U}^+(x_0)$ 时, $f(x_0) > f(x)$. 这样一来, 函数在整个邻域 $U(x_0)$ 中严格递减, 于是 x_0 不是极值点.

b) 首先, 像在 a) 中那样, 由于 $f(x)$ 在 $\overset{\circ}{U}^-(x_0)$ 递减和 f 在 x_0 连续, 我们断定, 当 $x \in \overset{\circ}{U}^-(x_0)$ 时, 有 $f(x) > f(x_0)$. 从 f 在 $\overset{\circ}{U}^+(x_0)$ 递增和 f 在 x_0 连续断定, 当 $x \in \overset{\circ}{U}^+(x_0)$ 时, $f(x_0) < f(x)$. 于是函数 f 在 x_0 有严格局部极小值.

断言 c) 和 d) 可类似地证明. ▶

命题 4 (用高阶导数表达的极值的充分条件) 设函数 $f: U(x_0) \to \mathbb{R}$ 定义在点 x_0 的邻域 $U(x_0)$ 中, 在 x_0 有直到 n 阶导数 $(n > 1)$.

如果 $f'(x_0) = \cdots = f^{(n-1)}(x_0) = 0$ 且 $f^{(n)}(x_0) \neq 0$, 则当 n 为奇数时, 在 x_0 处 f 无极值, 而当 n 为偶数时, f 有极值, 此时若 $f^{(n)}(x_0) > 0$, 则有严格局部极小值, 而若 $f^{(n)}(x_0) < 0$, 则有严格局部极大值.

◀ 我们将利用局部泰勒公式

$$f(x) - f(x_0) = \frac{1}{n!}f^{(n)}(x_0)(x - x_0)^n + \alpha(x)(x - x_0)^n, \tag{1}$$

其中当 $x \to x_0$ 时, $\alpha(x) \to 0$, 像在证明费马引理时那样进行讨论. 把 (1) 改写成

$$f(x) - f(x_0) = \left(\frac{1}{n!}f^{(n)}(x_0) + \alpha(x)\right)(x - x_0)^n. \tag{2}$$

因为 $f^{(n)}(x_0) \neq 0$, 而当 $x \to x_0$ 时, $\alpha(x) \to 0$, 所以, 当 x 充分接近 x_0 时, 和式 $f^{(n)}(x_0) + \alpha(x)$ 的值与 $f^{(n)}(x_0)$ 同号. 若 n 是奇数, 则经过 x_0 点时 $(x - x_0)^n$ 改变符号, 这时 (2) 式右端, 从而等式的左端都变号. 这就表明, 当 $n = 2k + 1$ 时没有极值.

若 n 为偶数, 则当 $x \neq x_0$ 时, $(x - x_0)^n > 0$, 因此在点 x_0 的小邻域内, 由等式 (2) 可见差 $f(x) - f(x_0)$ 的符号是与 $f^{(n)}(x_0)$ 的符号一样的. ▶

我们考察一些例子.

例 6 *几何光学中光的折射定律* (斯涅耳[①] 定律) 根据费马原理, 在任意两点之间, 光线所走的实际路径是耗费时间最少的路径.

从费马原理和两点间最短路径是以它们为端点的直线段这一事实推出, 在同一种均匀介质 (在每点和每个方向都有同样结构的介质) 中, 光是沿直线传播的.

现设有两种均匀介质且光从点 A_1 传到点 A_2, 如图 22 所示.

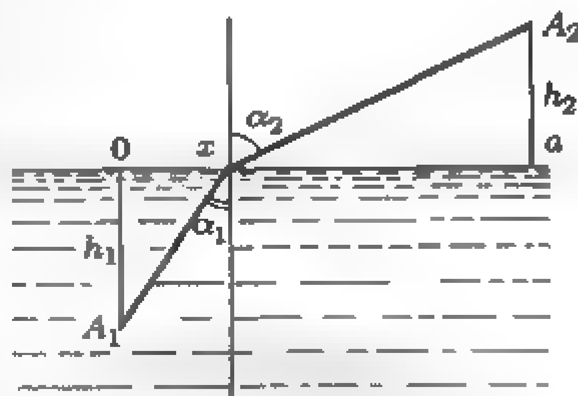


图 22

若 c_1、c_2 是光在这两种介质中的速度, 则通过所示路径的时间是:

$$t(x) = \frac{1}{c_1}\sqrt{h_1^2 + x^2} + \frac{1}{c_2}\sqrt{h_2^2 + (a - x)^2}.$$

我们来求函数 $t(x)$ 的极值:

$$t'(x) = \frac{1}{c_1}\frac{x}{\sqrt{h_1^2 + x^2}} - \frac{1}{c_2}\frac{a - x}{\sqrt{h_2^2 + (a - x)^2}} = 0,$$

①斯涅耳(Snell 或 Snellius) (1580—1626)—— 荷兰天文学家和数学家.

按照图上的记号, 此式给出

$$\frac{1}{c_1}\sin\alpha_1=\frac{1}{c_2}\sin\alpha_2.$$

从物理的考虑或者从 $x\to\infty$ 时函数 $t(x)$ 无限增长的情况, 都可以明白, 使 $t'(x)=0$ 的点是连续函数 $t(x)$ 的绝对极小值点. 于是从费马原理推出了折射定律

$$\frac{\sin\alpha_1}{\sin\alpha_2}=\frac{c_1}{c_2}.$$

例 7 我们证明: 对于 $x>0$ 有

$$x^\alpha-\alpha x+\alpha-1\leqslant 0,\quad 当0<\alpha<1时, \tag{3}$$

$$x^\alpha-\alpha x+\alpha-1\geqslant 0,\quad 当\alpha<0或1<\alpha时. \tag{4}$$

◀ 对函数 $f(x)=x^\alpha-\alpha x+(\alpha-1)$ 求微分, 得到 $f'(x)=\alpha(x^{\alpha-1}-1)$ 以及当 $x=1$ 时, $f'(x)=0$. 如果 $0<\alpha<1$, 则当点从 1 处经过时, 导数从正值变到负值, 而若 $\alpha<0$ 或 $1<\alpha$, 则当点从 1 处经过时, 导数从负值变为正值. 在第一种情形, 在点 1 处有严格极大值, 而在第二种情形则有严格极小值 (并且根据 f 在区间 $0<x<1$ 上的单调性知道这个极值不仅仅是局部的). 然而 $f(1)=0$, 于是不等式 (3)、(4) 都成立. 甚至还同时证明了, 如果 $x\neq 1$ 的话, 两个不等式都是严格的. ▶

我们注意到, 如果以 $1+x$ 代替 x, 我们将发现, (3) 和 (4) 是我们熟悉的指数是自然数的伯努利不等式的推广 (第二章 §2; 也可参看这一节末尾的题目 2).

借助于初等的代数变换, 从所证的不等式可以得到一系列在分析学中很重要的经典的不等式. 我们来推导这些不等式.

a. 杨格①不等式 如果 $a>0$ 且 $b>0$, 而数 p、q 满足:$p\neq 0、1, q\neq 0、1, \frac{1}{p}+\frac{1}{q}=1$, 那么

$$a^{\frac{1}{p}}b^{\frac{1}{q}}\leqslant\frac{1}{p}a+\frac{1}{q}b,\quad 当p>1, \tag{5}$$

$$a^{\frac{1}{p}}b^{\frac{1}{q}}\geqslant\frac{1}{p}a+\frac{1}{q}b,\quad 当p<1, \tag{6}$$

且 (5)、(6) 中的等号仅当 $a=b$ 时成立.

◀ 只要在 (3) 和 (4) 中置 $x=\frac{a}{b}$ 和 $\alpha=\frac{1}{p}$, 并引用记号 $\frac{1}{q}=1-\frac{1}{p}$, 便可证明. ▶

①杨格 (Young)(1882—1946)—— 英国数学家.

b. 赫尔德[①]不等式 设 $x_i \geqslant 0, y_i \geqslant 0\ (i=1,\cdots,n)$ 且 $\frac{1}{p}+\frac{1}{q}=1$. 那么

$$\sum_{i=1}^{n} x_i y_i \leqslant \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^{n} y_i^q\right)^{\frac{1}{q}}, \text{当} p>1, \tag{7}$$

$$\sum_{i=1}^{n} x_i y_i \geqslant \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^{n} y_i^q\right)^{\frac{1}{q}}, \text{当} p<1, p\neq 0. \tag{8}$$

当 $p<0$ 时,(8) 中假定 $x_i>0\ (i=1,\cdots,n)$. (7)、(8) 中的等号仅当向量 $(x_1^p,\cdots,x_n^p)$ 和 $(y_1^q,\cdots,y_n^q)$ 共线时成立.

◀ 我们验证不等式 (7). 设 $X=\sum_{i=1}^{n} x_i^p>0, Y=\sum_{i=1}^{n} y_i^q>0$. 在 (5) 中置 $a=\frac{x_i^p}{X}, b=\frac{y_i^q}{Y}$, 得到

$$\frac{x_i y_i}{X^{\frac{1}{p}} Y^{\frac{1}{q}}} \leqslant \frac{1}{p}\frac{x_i^p}{X}+\frac{1}{q}\frac{y_i^q}{Y}.$$

把这些不等式关于 i 从 1 到 n 加起来, 得到

$$\frac{\sum_{i=1}^{n} x_i y_i}{X^{\frac{1}{p}} Y^{\frac{1}{q}}} \leqslant 1,$$

这与 (7) 等价.

类似地从 (6) 得到 (8). 由于 (5) 和 (6) 中的等号仅当 $a=b$ 时成立, 所以我们判定 (7) 和 (8) 中的等号仅当 $x_i^p=\lambda y_i^q$ 或者 $y_i^q=\lambda x_i^p$ 时成立. ▶

c. 闵可夫斯基[②] 不等式 设 $x_i \geqslant 0, y_i \geqslant 0\ (i=1,\cdots,n)$. 那么

$$\left(\sum_{i=1}^{n}(x_i+y_i)^p\right)^{\frac{1}{p}} \leqslant \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} y_i^p\right)^{\frac{1}{p}}, \text{当} p>1, \tag{9}$$

$$\left(\sum_{i=1}^{n}(x_i+y_i)^p\right)^{\frac{1}{p}} \geqslant \left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} + \left(\sum_{i=1}^{n} y_i^p\right)^{\frac{1}{p}}, \text{当} p<1, p\neq 0. \tag{10}$$

◀ 对恒等式

$$\sum_{i=1}^{n}(x_i+y_i)^p \equiv \sum_{i=1}^{n} x_i(x_i+y_i)^{p-1} + \sum_{i=1}^{n} y_i(x_i+y_i)^{p-1}$$

右端的两项用赫尔德不等式.

① 赫尔德 (Hölder)(1859—1937)—— 德国数学家.

② 闵可夫斯基 (Minkowski)(1864—1909)—— 德国数学家, 提出了狭义相对论的恰当的数学模型 (不定度量空间).

那么左端，根据不等式 (7)、(8)，相应地被量

$$\left(\sum_{i=1}^{n} x_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n}(x_i+y_i)^p\right)^{\frac{1}{q}} + \left(\sum_{i=1}^{n} y_i^p\right)^{\frac{1}{p}} \left(\sum_{i=1}^{n}(x_i+y_i)^p\right)^{\frac{1}{q}}$$

从上方或从下方控制.

用 $\left(\sum_{i=1}^{n}(x_i+y_i)^p\right)^{\frac{1}{q}}$ 去除所得的不等式之后就得到 (9) 和 (10).

知道了在赫尔德不等式中等号成立的条件，我们就能验证在闵可夫斯基不等式中等号仅当向量 $(x_1, \cdots, x_n)$、$(y_1, \cdots, y_n)$ 共线时成立. ►

当 $n=3$、$p=2$ 时，闵可夫斯基不等式 (9) 显然就是三维欧几里得空间中的三角形不等式.

例 8 再考虑一个用高阶导数求局部极值的最简单的例子. 设 $f(x)=\sin x$. 因为 $f'(x)=\cos x$ 和 $f''(x)=-\sin x$，所以一切使 $f'(x)=\cos x=0$ 的点都是函数 $\sin x$ 的局部极值点，因为在这些点上 $f''(x)=-\sin x\neq 0$. 并且若 $\sin x>0$ 则 $f''(x)<0$; 而若 $\sin x<0$ 则 $f''(x)>0$. 于是使 $\cos x=0$ 且 $\sin x>0$ 的点是函数 $\sin x$ 的局部极大值点，而使 $\cos x=0$ 且 $\sin x<0$ 的点是它的局部极小值点 (这些当然都是早知如此的).

3. 函数凸的条件

定义 1 定义在开区间 $]a,b[\subset\mathbb{R}$ 上的函数 $f:]a,b[\to\mathbb{R}$ 叫 $]a,b[$ 上的凸函数，如果对于任意的点 x_1、$x_2\in]a,b[$ 和任意的数 $\alpha_1\geqslant 0$、$\alpha_2\geqslant 0$ 使 $\alpha_1+\alpha_2=1$ 者，成立不等式

$$f(\alpha_1x_1+\alpha_2x_2)\leqslant \alpha_1f(x_1)+\alpha_2f(x_2). \tag{11}$$

若当 $x_1\neq x_2$ 且 $\alpha_1\cdot\alpha_2\neq 0$ 时，这个不等式总是严格的，则函数叫做开区间 $]a,b[$ 上的严格凸函数.

从几何上来说，函数 $f:]a,b[\to\mathbb{R}$ 凸性的条件表示函数图像的任何一段弧上的点都位于这段弧所张的弦的下面 (图 23).

事实上，(11) 的左边是函数 $f(x)$ 在点 $x=\alpha_1x_1+\alpha_2x_2\in[x_1,x_2]$ 处的值，而右边是一个线性函数在同一点处的值，此线性函数的图像 (直线) 通过点 $(x_1,f(x_1))$、$(x_2,f(x_2))$.

关系式 (11) 还可加以更好的解释，就是说，平面上位于函数图像上方的点的集合

$$E=\{(x,y)\in\mathbb{R}^2|x\in]a,b[, f(x)<y\}$$

是凸集，"凸" 函数一语本身就是由此得来的.

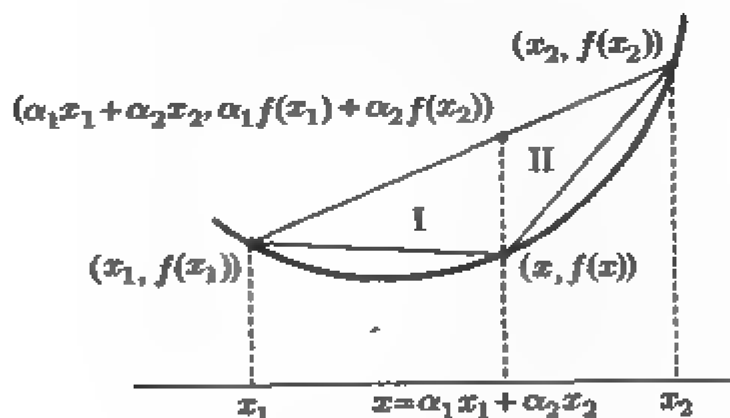


图 23

定义 2 若对于函数 $f:]a,b[\to\mathbb{R}$ 在 (11) 中成立着相反的不等式, 则说函数是开区间 $]a,b[$ 上的凹函数, 也常常说它是这个区间上的上凸函数以区别于凸函数, 而把凸函数叫做开区间 $]a,b[$ 上的下凸函数.

由于下凸函数和上凸函数的一切进一步的构造都是完全一样的, 所以我们只限于讨论 (下) 凸函数.

首先我们给出不等式 (11) 的另一个更适合于我们的目的的形式.

从关系式 $x=\alpha_1x_1+\alpha_2x_2, \alpha_1+\alpha_2=1$ 有 $\alpha_1=\dfrac{x_2-x}{x_2-x_1}$、$\alpha_2=\dfrac{x-x_1}{x_2-x_1}$, 因此 (11) 可改写成

$$f(x)\leqslant\frac{x_2-x}{x_2-x_1}f(x_1)+\frac{x-x_1}{x_2-x_1}f(x_2).$$

考虑到 $x_1\leqslant x\leqslant x_2$ 和 $x_1<x_2$, 乘以 x_2-x_1 之后得到

$$(x_2-x)f(x_1)+(x_1-x_2)f(x)+(x-x_1)f(x_2)\geqslant 0.$$

注意到 $x_2-x_1=x_2-x+x-x_1$, 从最后这个不等式经初等变换后得到: 当 $x_1<x<x_2$、x_1、$x_2\in]a,b[$ 时,

$$\frac{f(x)-f(x_1)}{x-x_1}\leqslant\frac{f(x_2)-f(x)}{x_2-x}. \tag{12}$$

不等式 (12) 是开区间 $]a,b[$ 上函数凸性定义的另一种形式. (12) 式在几何上表示联结点 $(x_1,f(x_1))$ 和 $(x,f(x))$ 的弦 I 的斜率不超过 (在严格凸的情形下是小于) 联结点 $(x,f(x))$ 和 $(x_2,f(x_2))$ 的弦 II 的斜率 (见图 23).

现在假定函数 $f:]a,b[\to\mathbb{R}$ 在 $]a,b[$ 上可微. 那么, 在 (12) 中先后令 $x\to x_1$ 和 $x\to x_2$, 就得到

$$f'(x_1)\leqslant\frac{f(x_2)-f(x_1)}{x_2-x_1}\leqslant f'(x_2),$$

这就确定了函数 f 的导数的单调性.

对于严格凸函数, 考虑到导数的单调性, 利用拉格朗日定理求得

$$f'(x_1)\leqslant f'(\xi_1)=\frac{f(x)-f(x_1)}{x-x_1}<\frac{f(x_2)-f(x)}{x_2-x}=f'(\xi_2)\leqslant f'(x_2),$$

当 $x_1 < \xi_1 < x < \xi_2 < x_2$ 时成立, 因此严格凸性蕴含着导数的严格单调性.

于是, 若可微函数 f 在开区间 $]a,b[$ 上是凸的, 则 f' 在 $]a,b[$ 上不减, 而在 f 严格凸的情形, 它的导数 f' 在 $]a,b[$ 上递增.

其实, 对于可微函数的凸性, 这不仅是必要条件, 也是充分条件.

事实上, 对于 $a < x_1 < x < x_2 < b$, 按拉格朗日定理

$$\frac{f(x)-f(x_1)}{x-x_1} = f'(\xi_1), \frac{f(x_2)-f(x)}{x_2-x} = f'(\xi_2),$$

其中 $x_1 < \xi_1 < x < \xi_2 < x_2$, 于是若 $f'(\xi_1) \leqslant f'(\xi_2)$ 则凸性的条件 (12) 成立 (而若 $f'(\xi_1) < f'(\xi_2)$, 则严格凸性的条件成立).

这样我们证明了下面的

命题 5 要使在开区间 $]a,b[$ 上可微的函数 $f:]a,b[\to \mathbb{R}$ 在 $]a,b[$ 上是 (下) 凸的, 必须且只需它的导数 f' 在 $]a,b[$ 上不减. 同时, f 的严格凸性对应着 f' 的严格递增.

对比命题 5 和命题 1, 得到

推论 要使在开区间 $]a,b[$ 上有二阶导数的函数 $f:]a,b[\to \mathbb{R}$ 在这个区间上是 (下) 凸的, 必须且只需在 $]a,b[$ 上有 $f''(x) \geqslant 0$. 如果 $f''(x) > 0$ 在 $]a,b[$ 上成立的话, 那么这已充分保障函数 $f:]a,b[\to \mathbb{R}$ 是严格凸的.

现在我们能够解释清楚, 例如, 为什么最简单的初等函数的图像总具有某种凸性.

例 9 研究函数 $f(x) = x^\alpha$ 在集合 $x > 0$ 上的凸性. 由于 $f''(x) = \alpha(\alpha-1)x^{\alpha-2}$, 那么, 当 $\alpha < 0$ 或 $\alpha > 1$ 时, $f''(x) > 0$, 所以对于幂指数 α 的这样的值, 幂函数 x^α 严格 (下) 凸. 当 $0 < \alpha < 1$ 时有 $f''(x) < 0$, 因此, 对于这样的幂指数, 它严格上凸. 例如, 我们总把抛物线 $f(x) = x^2$ 画得向下凸. 剩下的情形 $\alpha = 0$ 和 $\alpha = 1$ 是平庸的:$x^0 \equiv 1, x^1 = x$. 在这两种情况, 函数的图像都是射线 (见第 304 页上的图 30)

例 10 设 $f(x) = a^x, 0 < a, a \neq 1$. 由于 $f''(x) = a^x \ln^2 a > 0$, 所以, 对于任意许可的底数 a, 指数函数 a^x 都是在 $\mathbb{R}$ 上严格 (下) 凸的 (见第 304 页上的图 24).

例 11 对于函数 $f(x) = \log_a x$ 有 $f''(x) = -\dfrac{1}{x^2 \ln a}$, 因此, 当 $0 < a < 1$ 时函数严格 (下) 凸, 而当 $1 < a$ 时严格上凸 (见第 304 页上的图 25).

例 12 我们来研究函数 $f(x) = \sin x$ 的凸性 (见第 304 页上的图 26).

因为 $f''(x) = -\sin x$, 所以, 在开区间 $\pi 2k < x < \pi(2k+1)$ 上 $f''(x) < 0$, 而在开区间 $\pi(2k-1) < x < \pi 2k$ 上 $f''(x) > 0$, 其中 $k \in \mathbf{Z}$. 由此, 作为例子推出, 函数 $\sin x$ 的图像在闭区间 $0 \leqslant x \leqslant \dfrac{\pi}{2}$ 上的弧除端点外处处位于它所张的弦的上方; 因此当 $0 < x < \dfrac{\pi}{2}$ 时, $\sin x > \dfrac{2}{\pi}x$.

现在我们还要指出凸函数的一个特征，它在几何上等价于平面上的凸区域位于它的边界的切线的一侧这样一个性质.

命题 6 在开区间 $]a,b[$ 上可微的函数 $f:]a,b[\to\mathbb{R}$ 在 $]a,b[$ 是 (下) 凸的当且仅当函数的图像的一切点都不位于此图像的任何一条切线的下方. 同时，要使函数是严格凸的，必须且只需图像上所有的点除了切点本身以外都严格地位于这条切线的上方.

◀ **必要性** 设 $x_0\in]a,b[$. 图像在点 $(x_0,f(x_0))$ 处的切线的方程是

$$y=f(x_0)+f'(x_0)(x-x_0),$$

因此

$$\begin{aligned}f(x)-y(x)&=f(x)-f(x_0)-f'(x_0)(x-x_0)\\&=(f'(\xi)-f'(x_0))(x-x_0),\end{aligned}$$

其中 ξ 是 x 和 x_0 之间的点. 因为 f 是凸的，函数 $f'(x)$ 在 $]a,b[$ 上不减，于是差 $f'(\xi)-f'(x_0)$ 与差 $x-x_0$ 同号，因此在任意点 $x\in]a,b[$ 处，$f(x)-y(x)\geqslant 0$. 如果 f 是严格凸的，那么 f' 在 $]a,b[$ 上严格递增，这就表明，当 $x\in]a,b[$ 且 $x\neq x_0$ 时，$f(x)-y(x)>0$.

充分性 若对于任意的点 x、$x_0\in]a,b[$,

$$f(x)-y(x)=f(x)-f(x_0)-f'(x_0)(x-x_0)\geqslant 0,\tag{13}$$

则

$$\begin{aligned}&\frac{f(x)-f(x_0)}{x-x_0}\leqslant f'(x_0),\quad 当x<x_0时,\\&\frac{f(x)-f(x_0)}{x-x_0}\geqslant f'(x_0),\quad 当x_0<x时.\end{aligned}$$

于是，对于任意三点 x_1、x、$x_2\in]a,b[$，当 $x_1<x<x_2$ 时得到

$$\frac{f(x)-f(x_1)}{x-x_1}\leqslant\frac{f(x_2)-f(x)}{x_2-x},$$

同时,(13) 中的严格不等式蕴含着上式中的严格不等式. 我们看到，这个不等式与凸函数的定义的写法 (12) 是一样的. ▶

我们看几个例子.

例 13 函数 $f(x)=e^x$ 严格凸. 直线 $y=x+1$ 是这个函数的图像在点 $(0,1)$ 处的切线，因为 $f(0)=e^0=1$ 且 $f'(0)=e^x|_{x=0}=1$. 据命题 6 断定，对于任意的 $x\in\mathbb{R}$

$$e^x\geqslant 1+x,$$

且若 $x\neq 0$，则不等式是严格的.

例 14　类似地, 利用函数 $\ln x$ 的严格上凸性质, 可以验证当 $x>0$ 时, 成立着不等式

$$\ln x \leqslant x-1,$$

并且当 $x \neq 1$ 时, 这个不等式是严格的.

在画函数的图像时, 把图像的 "拐点" 挑出来常常是有益的.

定义 3　设 $f: U(x_0) \to \mathbb{R}$ 是在点 $x_0 \in \mathbb{R}$ 的邻域 $U(x_0)$ 中定义且可微的函数. 若在集

$$\overset{\circ}{U}{}^{-}(x_0)=\{x \in U(x_0) | x<x_0\}$$

上函数下 (上) 凸, 而在集

$$\overset{\circ}{U}{}^{+}(x_0)=\{x \in U(x_0) | x_0<x\}$$

上函数上 (下) 凸, 则图像的点 $(x_0, f(x_0))$ 叫做它的*拐点*.

于是, 在通过拐点时, 图像凸的方向发生变化, 特别地, 这表明, 在点 $(x_0, f(x_0))$ 处, 函数的图像从它在此点的切线的一侧转到另一侧.

对比命题 5 和命题 3, 容易看出拐点横坐标 x_0 的分析特征. 即可以说, 若 f 在点 x_0 二次可微, 那么, 由于 $f'(x)$ 在点 x_0 处有极值, 从而必须有 $f''(x_0)=0$.

若二阶导数 $f''(x)$ 在 $U(x_0)$ 中都有定义且在 $\overset{\circ}{U}{}^{-}(x_0)$ 中 $f''(x)$ 处处有同一符号, 但在 $\overset{\circ}{U}{}^{+}(x_0)$ 中处处有与它在 $\overset{\circ}{U}{}^{-}(x_0)$ 中相反的符号, 则这就足以使得 $f'(x)$ 在 $\overset{\circ}{U}{}^{-}(x_0)$ 中和在 $\overset{\circ}{U}{}^{+}(x_0)$ 中都是单调的, 但单调性质却不同. 那么根据命题 5, 在点 $(x_0, f(x_0))$ 处图像凸的方向发生变化, 即 $(x_0, f(x_0))$ 是拐点.

例 15　在例 12 中考察了函数 $f(x)=\sin x$, 求出了它的图像向上凸的区间和向下凸的区间. 现在我们证明图像上以 $x=\pi k$, $k \in \mathbb{Z}$ 为横坐标的点是拐点.

实际上, $f''(x)=-\sin x$; 当 $x=\pi k, k \in \mathbb{Z}$ 时, $f''(x)=0$. 此外, 当从这些点经过时 $f''(x)$ 变号, 这是拐点的充分条件 (见第 304 页的图 26).

例 16　不应该认为, 曲线在某点处从切线的一侧转到另一侧是判断此点为拐点的充分条件. 须知可以发生这样的情形, 在这点的无论是左邻域中还是右邻域中曲线都不保持确定的凸性. 例子是容易构造的, 只要把基于类似理由引入的例 5 做些改善即可.

设

$$f(x)=\begin{cases}2x^3+x^3 \sin \dfrac{1}{x^2}, & \text{当} x \neq 0 \text{时}, \\ 0, & \text{当} x=0 \text{时},\end{cases}$$

那么, 当 $0 \leqslant x$ 时, $x^3 \leqslant f(x) \leqslant 3x^3$, 而当 $x \leqslant 0$ 时, $3x^3 \leqslant f(x) \leqslant x^3$. 因此, 函数的图像在点 $x=0$ 处与横轴相切并在这点从下半平面转到上半平面. 同时, 函数 $f(x)$

的导数

$$f'(x)=\begin{cases}6x^2+3x^2\sin\dfrac{1}{x^2}-2\cos\dfrac{1}{x^2}, & 当x\neq 0时,\\ 0, & 当x=0时.\end{cases}$$

在点 $x=0$ 的无论哪个半邻域中都不单调.

最后我们重新回到凸函数的定义 (11) 并证明下面的

命题 7 (詹生[①]不等式) *若 $f:]a,b[\to\mathbb{R}$ 是凸函数, $x_1,\cdots,x_n$ 是开区间 $]a,b[$ 的点, $\alpha_1,\cdots,\alpha_n$ 是非负实数使 $\alpha_1+\cdots+\alpha_n=1$, 则成立不等式*

$$f(\alpha_1x_1+\cdots+\alpha_nx_n)\leqslant\alpha_1f(x_1)+\cdots+\alpha_nf(x_n). \tag{14}$$

◀ 当 $n=2$ 时,(14) 与凸函数的定义 (11) 一样.

我们来证, 若 (14) 对于 $n=m-1$ 成立, 则它对于 $n=m$ 也成立.

为确定起见, 设在 $\alpha_1,\cdots,\alpha_n$ 这组数中 $\alpha_n\neq 0$. 那时,

$$\beta=\alpha_2+\cdots+\alpha_n>0 \text{ 且 } \frac{\alpha_2}{\beta}+\cdots+\frac{\alpha_n}{\beta}=1.$$

用函数的凸性, 由于

$$\alpha_1+\beta=1 \text{ 和 } \left(\frac{\alpha_2}{\beta}x_2+\cdots+\frac{\alpha_n}{\beta}x_n\right)\in]a,b[,$$

得

$$\begin{aligned}f(\alpha_1x_1+\cdots+\alpha_nx_n)&=f\left(\alpha_1x_1+\beta\left(\frac{\alpha_2}{\beta}x_2+\cdots+\frac{\alpha_n}{\beta}x_n\right)\right)\\&\leqslant\alpha_1f(x_1)+\beta f\left(\frac{\alpha_2}{\beta}x_2+\cdots+\frac{\alpha_n}{\beta}x_n\right).\end{aligned}$$

其次, 按归纳假定

$$f\left(\frac{\alpha_2}{\beta}x_2+\cdots+\frac{\alpha_n}{\beta}x_n\right)\leqslant\frac{\alpha_2}{\beta}f(x_2)+\cdots+\frac{\alpha_n}{\beta}f(x_n).$$

因此,

$$\begin{aligned}f(\alpha_1x_1+\cdots+\alpha_nx_n)&\leqslant\alpha_1f(x_1)+\beta f\left(\frac{\alpha_2}{\beta}x_2+\cdots+\frac{\alpha_n}{\beta}x_n\right)\\&\leqslant\alpha_1f(x_1)+\alpha_2f(x_2)+\cdots+\alpha_nf(x_n).\end{aligned}$$

根据归纳法原理断定, (14) 对于任意的 $n\in\mathbb{N}$ 成立.(对于 $n=1$ (14) 是平庸的.) ▶

我们指出, 如同在证明中见到的, 严格凸性对应于严格詹生不等式, 亦即, 若数 $\alpha_1,\cdots,\alpha_n$ 异于零, 则 (14) 中的等号当且仅当 $x_1=\cdots=x_n$ 时成立.

---

①詹生 (Jensen)(1859—1925) —— 丹麦数学家.

当然, 对于上凸函数得到的是与不等式 (14) 相反的不等式

$$f(\alpha_1 x_1+\cdots+\alpha_n x_n) \geqslant \alpha_1 f(x_1)+\cdots+\alpha_n f(x_n). \tag{15}$$

例 17 函数 $f(x)=\ln x$ 在正数集上严格上凸, 因此根据 (15), 对于 $x_i \geqslant 0, \alpha_i \geqslant 0\ (i=1,\cdots,n)$ 和 $\sum\limits_{i=1}^{n}\alpha_i=1$,

$$\alpha_1 \ln x_1+\cdots+\alpha_n \ln x_n \leqslant \ln(\alpha_1 x_1+\cdots+\alpha_n x_n)$$

或

$$x_1^{\alpha_1}\cdot\cdots\cdot x_n^{\alpha_n} \leqslant \alpha_1 x_1+\cdots+\alpha_n x_n. \tag{16}$$

特别地, 若 $\alpha_1=\cdots=\alpha_n=\dfrac{1}{n}$ 则得到 n 个非负实数的几何平均与算术平均之间的经典的不等式

$$\sqrt[n]{x_1\cdot\cdots\cdot x_n} \leqslant \frac{x_1+\cdots+x_n}{n}. \tag{17}$$

在 (17) 中, 像前面曾指出的那样, 等号仅当 $x_1=x_2=\cdots=x_n$ 时成立. 如若在 (16) 中置 $n=2, \alpha_1=\dfrac{1}{p}, \alpha_2=\dfrac{1}{q}, x_1=a, x_2=b$, 则将再次得到我们已经知道的不等式 (5).

例 18 设 $f(x)=x^p, x \geqslant 0, p>1$. 由于这样的函数是凸的, 所以有

$$\left(\sum_{i=1}^{n}\alpha_i x_i\right)^p \leqslant \sum_{i=1}^{n}\alpha_i x_i^p.$$

于此置

$$q=\frac{p}{p-1}, \alpha_i=\frac{b_i^q}{\sum\limits_{k=1}^{n} b_k^q}, x_i=\frac{a_i \sum\limits_{k=1}^{n} b_k^q}{b_i^{\frac{1}{p-1}}},$$

就再次得到赫尔德不等式 (7)

$$\sum_{i=1}^{n} a_i b_i \leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{\frac{1}{p}} \cdot \left(\sum_{i=1}^{n} b_i^q\right)^{\frac{1}{q}},$$

其中 $\dfrac{1}{p}+\dfrac{1}{q}=1$ 且 $p>1$.

当 $p<1$ 时, 函数 $f(x)=x^p$ 是上凸的, 因此, 经类似的讨论就可得到另一个赫尔德不等式 (8).

4. 洛必达法则 现在我们来研究求函数之比的极限的一个特殊的, 然而有时是很有效的方法, 这就是著名的洛必达[①]法则.

命题 8 (洛必达法则) 设函数 $f:]a,b[\to\mathbb{R}$ 和 $g:]a,b[\to\mathbb{R}$ 在开区间 $]a,b[$ 上可微 $(-\infty\leqslant a<b\leqslant+\infty)$, 并且在 $]a,b[$ 上 $g'(x)\neq 0$ 且

$$\frac{f'(x)}{g'(x)}\to A, \text{当} x\to a+0(-\infty\leqslant A\leqslant+\infty).$$

那么, 只要下面两种情况有一种成立, 就有

$$\frac{f(x)}{g(x)}\to A, \text{当} x\to a+0.$$

这两种情况是

1° $(f(x)\to 0)\wedge(g(x)\to 0)$, 当 $x\to a+0$ 时.

2° $g(x)\to\infty$, 当 $x\to a+0$ 时.

当 $x\to b-0$ 时类似的结论也成立.

人们常简单地但不十分确切地这样来叙述洛必达法则:函数比的极限等于它们的导数的比的极限, 只要后者存在.

◀ 由于在 $]a,b[$ 上 $g'(x)\neq 0$, 根据罗尔定理可以断言, $g(x)$ 在 $]a,b[$ 上严格单调. 因此, 只要另取 b, 把它取得与 a 充分近 (如果有必要这样做的话), 可以认为, 在 $]a,b[$ 上成立 $g(x)\neq 0$. 对于 $x,y\in]a,b[$, 根据柯西定理, 存在点 $\xi\in]a,b[$, 它在 x 和 y 之间, 且有

$$\frac{f(x)-f(y)}{g(x)-g(y)}=\frac{f'(\xi)}{g'(\xi)}.$$

我们把这个等式改写成方便目前应用的形式

$$\frac{f(x)}{g(x)}=\frac{f(y)}{g(x)}+\frac{f'(\xi)}{g'(\xi)}\left[1-\frac{g(y)}{g(x)}\right].$$

当 $x\to a+0$ 时, 与 x 的变化相适应, 我们令 y 这样地趋于 $a+0$, 使满足

$$\frac{f(y)}{g(x)}\to 0 \quad \text{且} \quad \frac{g(y)}{g(x)}\to 0.$$

容易看出, 无论是假定的情况 1° 还是情况 2°, 这总是办得到的. 由于 ξ 在 x 和 y 之间, 所以, 与 x 和 y 一起, 也有 $\xi\to a+0$: 因此, 最后这个等式的右边趋于 A, 从而, 它的左边也趋于 A. ▶

例 19 $\lim\limits_{x\to 0}\dfrac{\sin x}{x}=\lim\limits_{x\to 0}\dfrac{\cos x}{1}=1.$

---

①洛必达 (L' Hospital)(1661 —1704) —— 法国数学家, 约翰 · 伯努利的高才生, 侯爵约翰 · 伯努利于 1691 1692 年为他写了第一本分析学教科书. 洛必达将书中关于微分学的部分以稍微不同的形式用自己的名义发表. 因此, 我们把 "洛必达法则" 归功于约翰 · 伯努利.

不应认为这是证明当 $x \to 0$ 时有 $\dfrac{\sin x}{x} \to 1$ 的一个新方法. 问题在于, 譬如说, 在导出关系式 $\sin' x = \cos x$ 时, 我们就已经应用了这里要计算的极限.

只有求得了导数的比的极限, 才有应用洛必达法则的可能. 这时, 不应该忘记验证条件 1° 或 2°. 下面的例子表明了这些条件的重要性.

例 20 设 $f(x) = \cos x, g(x) = \sin x$. 那么 $f'(x) = -\sin x, g'(x) = \cos x$, 从而, 当 $x \to +0$ 时, $\dfrac{f(x)}{g(x)} \to +\infty$, 同时当 $x \to +0$ 时 $\dfrac{f'(x)}{g'(x)} \to 0$.

例 21 $\lim\limits_{x\to+\infty} \dfrac{\ln x}{x^\alpha} = \lim\limits_{x\to+\infty} \dfrac{\frac{1}{x}}{\alpha x^{\alpha-1}} = \lim\limits_{x\to+\infty} \dfrac{1}{\alpha x^\alpha} = 0$ 当 $\alpha > 0$.

例 22 当 $a > 1$ 时

$$\lim_{x\to+\infty} \frac{x^\alpha}{a^x} = \lim_{x\to+\infty} \frac{\alpha x^{\alpha-1}}{a^x \ln a} = \cdots = \lim_{x\to+\infty} \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)x^{\alpha-n}}{a^x \ln^n a} = 0.$$

这是因为当 $n > \alpha$ 时, 若 $x \to +\infty$, 由于 $a > 1$, 则显然有 $\dfrac{x^{\alpha-n}}{a^x} \to 0$.

我们看到, 最后这串等式直到我们得到可求其极限的表达式为止, 其中的每个比式都具有所要求的性质.

5. 作函数的图像 为了直观地描述函数, 常常用到它的图像表示. 通常, 在讨论函数性状的定性问题时, 这种图像表示是很有用处的.

人们极少用图像来做精确的计算. 因此, 实践中重要的不是用图像对函数做精细的描画, 而是做出能正确反映函数性状基本特征的函数图像草图. 在这一段我们将考察在作函数图像草图时常遇到的一些一般方法.

a. 初等函数的图像 首先我们要注意基本初等函数的图像是什么样子, 自如地掌握这些图像对于以后是必须的 (图 24—30).

b. 作函数图像草图的例子 (不涉及微分学). 现在我们来看这样一些例子, 在这些例子中, 只要我们已经知道最简单的初等函数的图像和性质, 就可以容易地画出函数图像的草图来.

例 23 画函数

$$y = \log_{x^2-3x+2} 2$$

的图像和草图.

考虑到

$$y = \log_{x^2-3x+2} 2 = \frac{1}{\log_2(x^2-3x+2)} = \frac{1}{\log_2[(x-1)(x-2)]},$$

先作二次三项式 $y_1 = x^2 - 3x + 2$ 的图像, 而后作 $y_2 = \log_2 y_1(x)$ 的图像, 最后作

$y = \dfrac{1}{y_2(x)}$ 的图像 (图 31).

函数图像的这个形式也可用另外方法去"猜测": 弄清函数

$$\log_{x^2-3x+2} 2 = (\log_2(x^2-3x+2))^{-1}$$

的定义域, 找出函数在接近定义域的边界点时以及在以定义域的边界点为端点的区间上的性状, 参照已求得的函数在区间端点附近的性状画出"平滑的曲线".

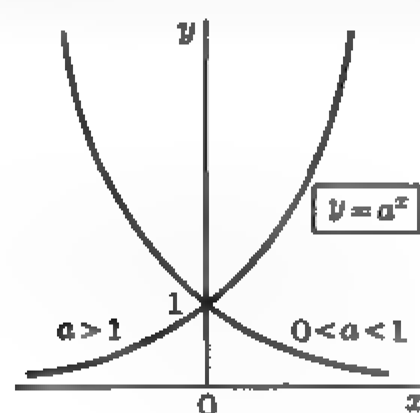


图 24

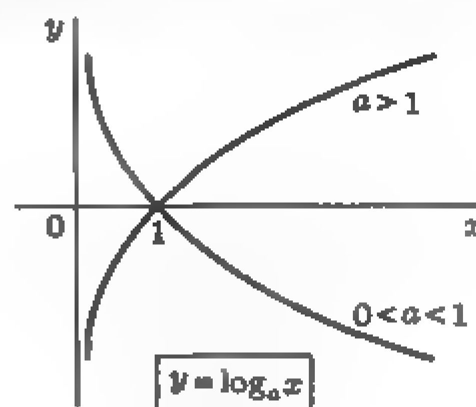


图 25

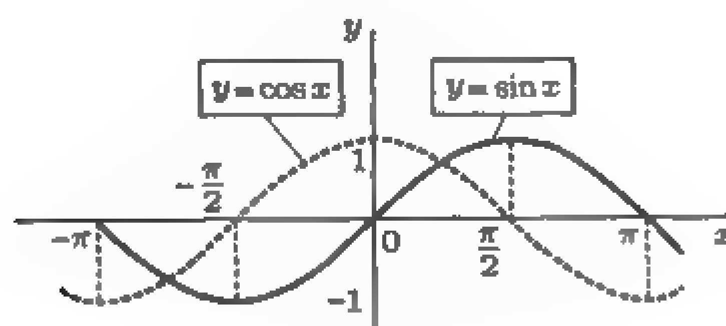


图 26

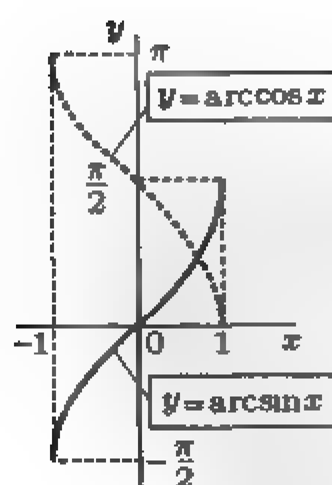


图 27

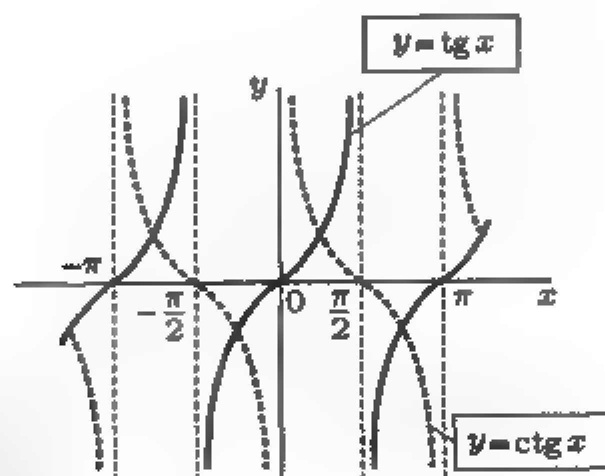


图 28

例 24 函数

$$y = \sin x^2$$

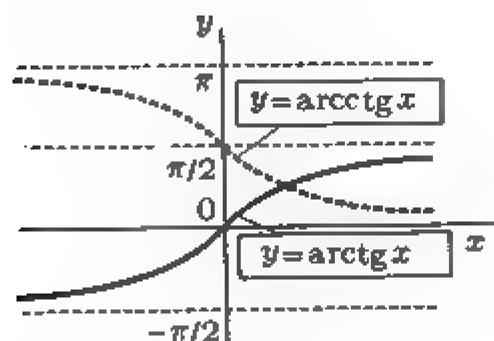


图 29

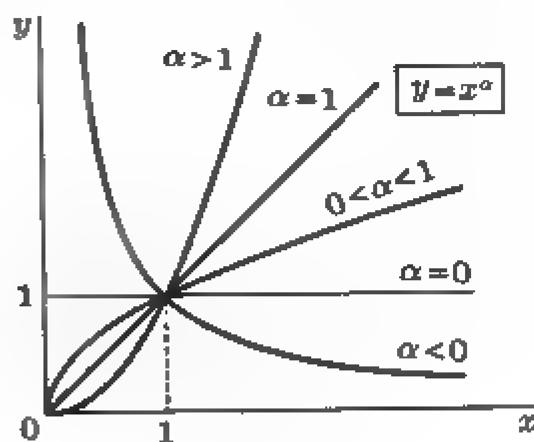


图 30

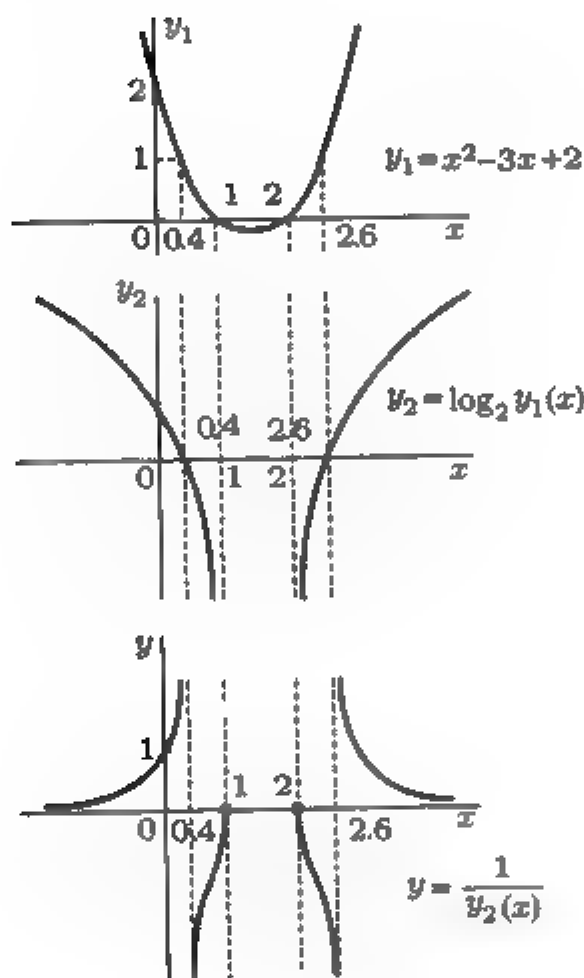


图 31

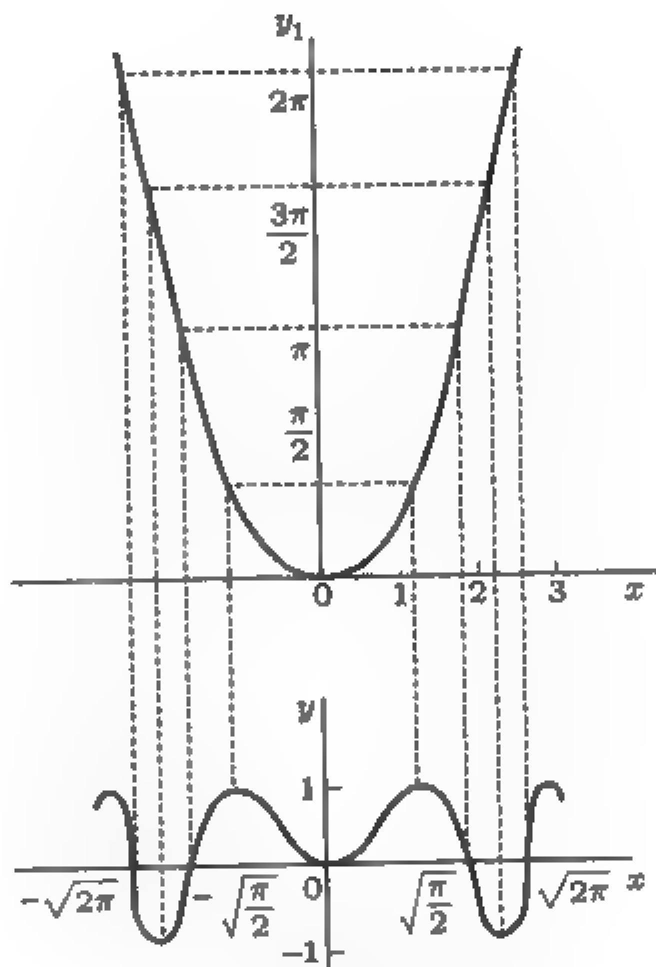


图 32

图像的草图见于图 32.

我们已经按照这个函数的一些特征性的点画出了它的图像. 这些特殊的点是使 $\sin x^2=-1$, $\sin x^2=0$ 及 $\sin x^2=1$ 的点. 在两个相邻的这种类型的点之间函数是单调的. 图像在点 $x=0, y=0$ 的邻域内的形状被这样的事实决定: 当 $x\to 0$ 时, $\sin x^2\sim x^2$. 此外, 注意到函数是偶函数是有益的.

由于我们始终只讲草图而不是精确作函数的图像, 所以为简单起见, 我们约定以后认为凡要求 "作函数的图像", 对于我们来说总等价于要求 "作函数图像的草图".

例 25 作函数

$$y = x + \operatorname{arctg}(x^3 - 1)$$

的图像 (图 33). 当 $x \to -\infty$ 时, 图像很接近直线 $y = x - \frac{\pi}{2}$, 而当 $x \to +\infty$ 时很接近直线 $y = x + \frac{\pi}{2}$.

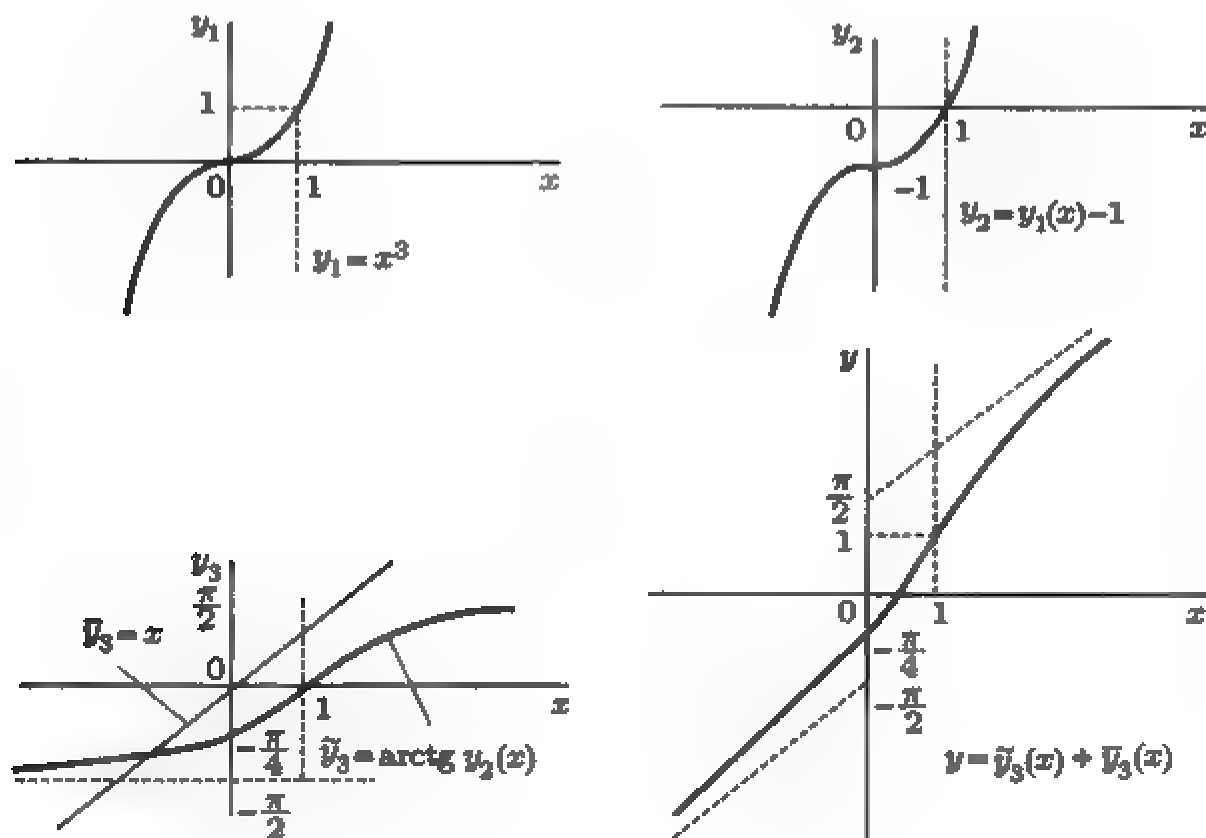


图 33

我们引入下述很有用的

定义 4 直线 $y = c_0 + c_1 x$ 叫做当 $x \to -\infty$(当 $x \to +\infty$) 时函数 $y = f(x)$ 的图像的渐近线, 如果当 $x \to -\infty$(当 $x \to +\infty$) 时 $f(x) - (c_0 + c_1 x) = o(1)$.

在例 25 中, 当 $x \to -\infty$ 时, 图像有渐近线 $y = x - \frac{\pi}{2}$, 而当 $x \to +\infty$ 时, 有渐近线 $y = x + \frac{\pi}{2}$.

若当 $x \to a - 0$(或当 $x \to a + 0$) 时, $|f(x)| \to \infty$, 那么, 很显然, 函数的图像这时随着 x 趋近于 a 而越来越贴近竖直线 $x = a$. 这条直线就叫做图像的竖直渐近线, 以区别于定义 4 中所说的那种永远倾斜渐近线.

因此, 例 23 中的函数图像 (见图 31) 有两条竖直的渐近线和一条 (对于 $x \to -\infty$ 和 $x \to +\infty$ 共有的) 水平渐近线.

从定义 4 显然可以推出

$$c_1 = \lim_{x \to -\infty} \frac{f(x)}{x},$$
$$c_0 = \lim_{x \to -\infty} (f(x) - c_1 x).$$

一般说来, 若当 $x \to -\infty$ 时, $f(x) - (c_0 + c_1x + \cdots + c_nx^n) = o(1)$, 则

$$c_n = \lim_{x\to-\infty} \frac{f(x)}{x^n},$$
$$c_{n-1} = \lim_{x\to-\infty} \frac{f(x) - c_nx^n}{x^{n-1}},$$
$$\cdots\cdots\cdots\cdots$$
$$c_0 = \lim_{x\to-\infty} f(x) - (c_1x + \cdots + c_nx^n).$$

我们对于 $x \to -\infty$ 的情形写出的这些关系式当然对于 $x \to +\infty$ 的情形也是对的, 它们都可以用于通过相应的代数多项式 $c_0 + c_1x + \cdots + c_nx^n$ 的图像描述函数 $f(x)$ 图像的渐近性状.

例 26 设 (ρ, φ) 是平面上的极坐标, 并设点在平面上这样运动: 在时刻 $t(t \geqslant 0)$

$$\rho = \rho(t) = 1 - e^{-t}\cos\frac{\pi}{2}t,$$
$$\varphi = \varphi(t) = 1 - e^{-t}\sin\frac{\pi}{2}t.$$

要求画出点的轨迹.

为此, 我们先画函数 $\rho(t)$ 和 $\varphi(t)$ 的图像 (图 34a、34b).

现在同时注视所作的这两个图像, 就能画出点的轨迹的一般形状 (图 34c).

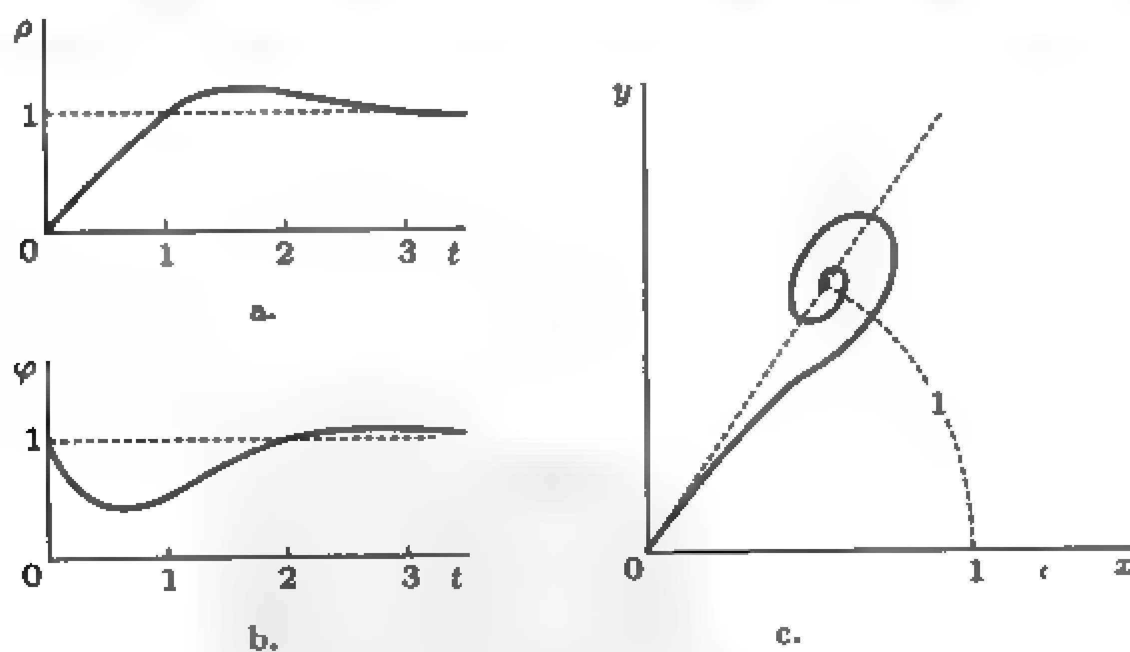


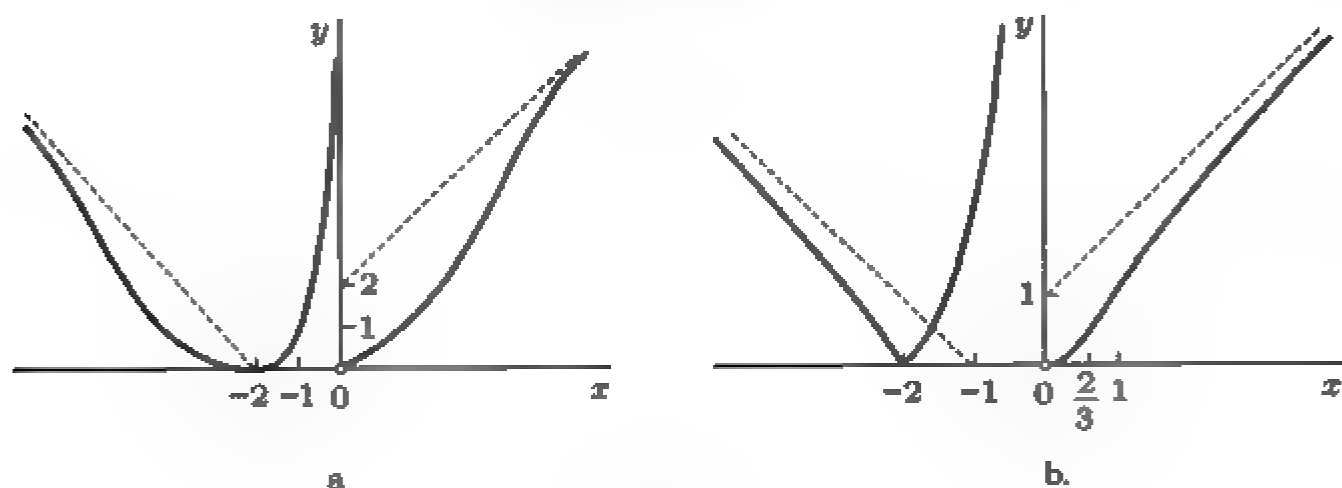
图 34

c. 微分学应用于作函数的图像 我们已经见到, 只根据最简单的一些考虑就可以大致画出很多函数的图像. 但是, 如果我们希望把草图画得更准确些, 那么, 当所研究的函数的导数不十分复杂的时候, 就可以使用微分学的工具. 我们通过例子来示范.

例 27 设 $f(x)=|x+2|e^{-\frac{1}{x}}$, 作函数 $y=f(x)$ 的图像. 函数 $f(x)$ 对于 $x\in\mathbb{R}\backslash 0$ 有定义. 因为当 $x\to\infty$ 时, $e^{-\frac{1}{x}}\to 1$, 所以

$$|x+2|e^{-\frac{1}{x}}\sim\begin{cases}-(x+2), \text{当} x\to-\infty \text{时},\\ (x+2), \quad \text{当} x\to+\infty \text{时}.\end{cases}$$

还有, 当 $x\to -0$ 时, 显然 $|x+2|e^{-\frac{1}{x}}\to+\infty$, 而当 $x\to+0$ 时, $|x+2|e^{-\frac{1}{x}}\to+0$. 最后, 显然 $f(x)\geqslant 0$ 且 $f(-2)=0$. 根据这些观察已经可以画出图像的初步轮廓 (图 35a).



a b.

图 35

现在我们来认真地弄清楚所给的函数是否真的在区间 $]-\infty,-2]$, $[-2,0[$, $]0,+\infty[$ 上单调, 它是否真有所指出的渐近线以及函数图像的凸性特征画得是否正确.

由于

$$f'(x)=\begin{cases}-\dfrac{x^2+x+2}{x^2}e^{-\frac{1}{x}}, \text{若} x<-2,\\ \dfrac{x^2+x+2}{x^2}e^{-\frac{1}{x}}, \quad \text{若}-2<x \text{且} x\neq 0,\end{cases}$$

以及 $f'(x)\neq 0$, 那么, 可编制表 5.2.

表 5.2

区间	$]-\infty,-2[$	$]-2,0[$	$]0,+\infty[$
$f'(x)$ 的符号	$-$	$+$	$+$
$f(x)$ 的性状	$+\infty\searrow 0$	$0\nearrow+\infty$	$0\nearrow+\infty$

在导数不变号的区间, 如我们已知的那样, 函数具有相应的单调性. 表中最下面一行里的符号 $+\infty\searrow 0$ 表示由 $+\infty$ 单调递减到 0, 而符号 $0\nearrow+\infty$ 表示函数的值由 0 单调递增到 $+\infty$.

我们指出, 当 $x\to-2-0$ 时, $f'(x)\to -e^{\frac{1}{2}}$, 而当 $x\to-2+0$ 时, $f'(x)\to e^{\frac{1}{2}}$, 因此, 点 $(-2,0)$ 应该是图像的角点 (有如同函数 $|x|$ 的图像中那种类型的转折), 而

不是像我们画在图 35a 上那样通常的点. 另外, 当 $x \to +0$ 时, $f'(x) \to 0$, 因此, 图像应该从原点出来与横轴相切 (想一想 $f'(x)$ 的几何意义!).

现在来准确地分析函数当 $x \to -\infty$ 和 $x \to +\infty$ 时的渐近性.

由于当 $x \to \infty$ 时, $e^{-\frac{1}{x}} = 1 - \frac{1}{x} + o\left(\frac{1}{x}\right)$, 所以

$$|x+2|e^{-\frac{1}{x}} = \begin{cases} -x-1+o(1), & \text{当}x \to -\infty\text{时}, \\ x+1+o(1), & \text{当}x \to +\infty\text{时}. \end{cases}$$

这就表明, 图像的倾斜渐近线其实是 $y = -x-1$(当 $x \to -\infty$ 时) 和 $y = x+1$(当 $x \to +\infty$ 时).

根据这些材料, 已经可以制作相当可靠的图像草图了. 但是, 我们继续深入一步, 在计算出

$$f''(x) = \begin{cases} -\dfrac{2-3x}{x^4}e^{-\frac{1}{x}}, & \text{若}x < -2, \\ \dfrac{2-3x}{x^4}e^{-\frac{1}{x}}, & \text{若} -2 < x\text{且}x \neq 0, \end{cases}$$

后再找出图像具凸性的段落.

因为仅当 $x = \frac{2}{3}$ 时, $f''(x) = 0$, 所以得到下表:

表 5.3

区间	$]-\infty,-2[$	$]-2,0[$	$]0,\frac{2}{3}[$	$]\frac{2}{3},+\infty[$
$f''(x)$ 的符号	−	+	+	−
$f(x)$ 的凸性	上	下	下	上

因为当 $x = \frac{2}{3}$ 时, 我们的函数可微, 而在通过这点时 $f''(x)$ 变号, 所以点 $\left(\frac{2}{3}, f\left(\frac{2}{3}\right)\right)$ 是图像的拐点.

顺便说一句, 假使导数 $f'(x)$ 成为零, 则从 $f'(x)$ 的符号的表中, 就可以判断在相应的点处有没有极值. 在我们的情况下, $f'(x)$ 无论在哪点都不为零, 但在点 $x = -2$ 处, 函数有局部极小值: 它在这点处连续, 并且通过这点时 $f'(x)$ 从负号变到正号. 其实, 我们的函数在 $x = -2$ 有最小值这一点, 从表中对函数值在相应区间上的变化的描述, 当然还要考虑到 $f(-2) = 0$ 就已经看出来了.

现在可以画出所给函数的更精确的图像草图了 (见图 35b).

最后我们再考察一个

例 28 设 (x, y) 是平面上的笛卡儿坐标, 并设动点在每个时刻 $t(t \geqslant 0)$ 时有坐标

$$x = \frac{t}{1-t^2}, \quad y = \frac{t-2t^3}{1-t^2}.$$

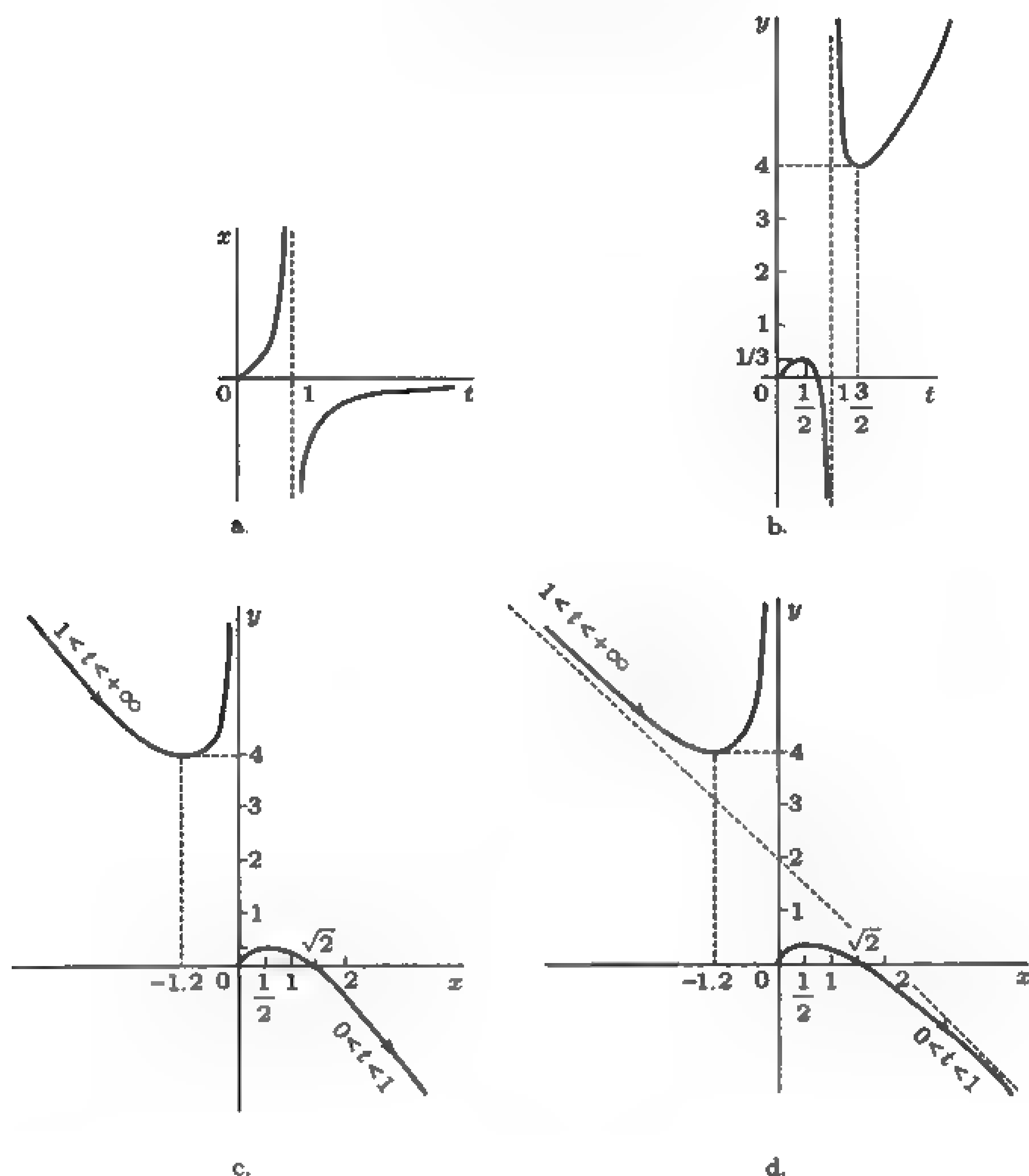


图 36

要求画出点运动的轨迹.

先画出所给的每个坐标函数 $x = x(t)$ 和 $y = y(t)$ 的图像的草图 (图 36a、36b)、这两个图像中的第二个更为有趣, 因此, 我们来弄清它的作法.

直接从 $y(t)$ 的解析表达式看出当 $t \to +0, t \to 1-0, t \to 1+0$ 时函数 $y = y(t)$ 的性状和当 $t \to +\infty$ 时函数的渐近式 $y(t) = 2t + o(1)$.

算出导数

$$\dot{y}(t) = \frac{1 - 5t^2 + 2t^4}{(1-t^2)^2}$$

之后, 求出它在区域 $t \geqslant 0$ 中的零点 $t_1 \approx 0.5$ 和 $t_2 \approx 1.5$.

造表:

表 5.4

区间	$]0, t_1[$	$]t_1, 1[$	$]1, t_2[$	$]t_2, +\infty[$
$\dot{y}(t)$ 的符号	+	−	−	+
$y(t)$ 的性状	$0 \nearrow y(t_1)$	$y(t_1) \searrow -\infty$	$+\infty \searrow y(t_2)$	$y(t_2) \nearrow +\infty$

然后求出使函数单调的区间和函数的局部极值 $y(t_1) \approx \frac{1}{3}$(极大) 和 $y(t_2) \approx 4$(极小).

现在兼顾 $x = x(t)$ 和 $y = y(t)$ 的图像, 画出点在平面上运动轨迹的草图 (见图 36c).

这个草图可以精确化. 例如, 可以弄清楚轨迹的渐近线.

由于 $\lim\limits_{t\to 1} \frac{y(t)}{x(t)} = -1$ 且 $\lim\limits_{t\to 1}(y(t) + x(t)) = 2$, 所以直线 $y = -x + 2$ 是对应于 $t \to 1$ 的轨迹两端的渐近线. 也很清楚, 直线 $x = 0$ 是相应于 $t \to +\infty$ 的那段轨迹的竖直渐近线.

再求出

$$y'_x = \frac{\dot{y}_t}{\dot{x}_t} = \frac{1 - 5t^2 + 2t^4}{1 + t^2}.$$

容易明白, 函数 $\frac{1 - 5u + 2u^2}{1 + u}$ 当 u 从 0 增至 1 时单调地从 1 递减到 -1, 而当 u 从 1 增至 $+\infty$ 时从 -1 递增到 $+\infty$.

从 y'_x 单调性的特征可以对轨迹在相应段落上的凸性的特征做出判断. 考虑到所说的这些, 现在可以做出点运动轨迹的如下更精确的草图 (见图 36d).

如果我们还考虑 $t < 0$ 时的轨迹的话, 那么, 从函数 $x(t)$ 和 $y(t)$ 的奇性推出, 对于平面 (x, y) 上已作好的曲线还要添上与它中心对称的曲线.

关于制作用解析式子给出的函数的图像的程序, 我们做个总结, 提出如下最一般的方案:

1° 指出函数的定义域.

2° 记下函数的明显特性 (例如, 偶性、奇性、周期性, 其图像经过最简单的坐标变换与已知函数图像重合).

3° 查明函数在趋向定义域的边界点时的渐近性状, 特别地, 如果有渐近线的话, 把渐近线求出来.

4° 求出函数单调的区间并标出函数的局部极值点.

5° 明确表出图像的凸性特征并标出拐点.

6° 标出图像的特殊点, 尤其是与坐标轴的交点, 如果这样的点存在且能算出的话.

练　习

1. 设 $x=(x_1,\cdots,x_n),\alpha=(\alpha_1,\cdots,\alpha_n)$, 且对于 $i=1,2,\cdots,n,x_i>0,\alpha_i>0$, 还有 $\sum\limits_{i=1}^{n}\alpha_i=1$. 对于任意的数 $t\neq 0$, 我们考察数 $x_1,x_2,\cdots,x_n$ 以 $\alpha_1,\cdots,\alpha_n$ 加权的 t 阶平均

$$M_t(x,\alpha)=\left(\sum_{i=1}^{n}\alpha_i x_i^t\right)^{\frac{1}{t}}.$$

特别地, 当 $\alpha_1=\cdots=\alpha_n=\dfrac{1}{n}$ 且 $t=-1,1,2$ 时, 分别得到调和平均、算术平均及二阶平均.

证明:

a) $\lim\limits_{t\to 0}M_t(x,\alpha)=x_1^{\alpha_1}\cdot\cdots\cdot x_n^{\alpha_n}$,
即几何平均可以作为极限而得到.

b) $\lim\limits_{t\to+\infty}M_t(x,\alpha)=\max\limits_{1\leqslant i\leqslant n}x_i$.

c) $\lim\limits_{t\to-\infty}M_t(x,\alpha)=\min\limits_{1\leqslant i\leqslant n}x_i$.

d) $M_t(x,\alpha)$ 在 $\mathbb{R}$ 上是 t 的非减函数 (M_0 由 a) 中极限定义), 并且若 $n>1$, 所有的 x_i 不同为一数, 则 $M_t(x,\alpha)$ 是严格递增的.

2. 证明:

$$|1+x|^p\geqslant 1+px+c_p\varphi_p(x),$$

其中 c_p 是只依赖于 p 的正的常数, 而当 $1<p\leqslant 2$ 时, $\varphi_p(x)=\begin{cases}|x|^2, 当|x|\leqslant 1时,\\ |x|^p, 当|x|>1时.\end{cases}$

当 $2<p$ 时, 在 $\mathbb{R}$ 上

$$\varphi_p(x)=|x|^p.$$

3. 验证: 当 $0<|x|<\dfrac{\pi}{2}$ 时, $\cos x<\left(\dfrac{\sin x}{x}\right)^3$.

4. 研究函数 $f(x)$ 并做出它的图像:

a) $f(x)=\operatorname{arctg}\log_2\cos\left(\pi x+\dfrac{\pi}{4}\right)$;

b) $f(x)=\arccos\left(\dfrac{3}{2}-\sin x\right)$;

c) $f(x)=\sqrt[3]{x(x+3)^2}$;

d) 做出在极坐标系中由方程 $\varphi=\dfrac{\rho}{\rho^2+1},\rho\geqslant 0$ 给定的曲线, 并标出它的渐近线.

e) 指出在知道了函数 $y=f(x)$ 的图像之后怎样得到函数 $f(x)+B$、$Af(x)$、$f(x+b)$、$f(ax)$, 特别是 $-f(x)$ 和 $f(-x)$ 的图像.

5. 证明: 如果 $f\in C(]a,b[)$ 且对于任意的点 x_1、$x_2\in]a,b[$ 成立不等式

$$f\left(\frac{x_1+x_2}{2}\right)\leqslant\frac{f(x_1)+f(x_2)}{2},$$

那么, 函数 f 在 $]a,b[$ 上是凸的.

6. 证明:

a) 若凸函数 $f:\mathbb{R}\to\mathbb{R}$ 有界, 则它是常值函数.

b) 若对于凸函数 $f:\mathbb{R}\to\mathbb{R}$,

$$\lim_{x\to-\infty}\frac{f(x)}{x}=\lim_{x\to+\infty}\frac{f(x)}{x}=0,$$

则 f 是常值函数.

c) 对于定义在区间 $a<x<+\infty$(或 $-\infty<x<a$) 上的任意的凸函数 f, 比值 $\frac{f(x)}{x}$ 当 x 沿着函数的定义域趋于无穷时趋于有限极限或趋于无穷.

7. 证明: 若 $f:]a,b[\to\mathbb{R}$ 是凸函数, 则

a) 在任意点 $x\in]a,b[$ 处, 它有左导数f'_- 和右导数f'_+:

$$f'_-(x)=\lim_{h\to-0}\frac{f(x+h)-f(x)}{h},f'_+(x)=\lim_{h\to+0}\frac{f(x+h)-f(x)}{h},$$

并且 $f'_-(x)\leqslant f'_+(x)$.

b) 当 x_1、$x_2\in]a,b[$ 且 $x_1<x_2$ 时成立不等式 $f'_+(x_1)\leqslant f'_-(x_2)$.

c) $f(x)$ 的图像的角点 (使 $f'_-(x)\neq f'_+(x)$ 的点) 是至多可数集.

8. 定义在区间 $I\subset\mathbb{R}$ 上的函数 $f:I\to\mathbb{R}$ 的勒让德[①]变换指的是函数

$$f^*(t)=\sup_{x\in I}(tx-f(x)).$$

证明:

a) 使 $f^*(t)\in\mathbb{R}$(即 $f^*(t)\neq\infty$) 的值 $t\in\mathbb{R}$ 的集合 I^* 或者是空集, 或者是单点集, 或者是区间, 而在最后一种情况, 函数 $f^*(t)$ 在 I^* 上是凸的.

b) 若 f 是凸函数, 则 $I^*\neq\varnothing$ 且当$f^*\in C(I^*)$时,

$$(f^*)^*(x)=\sup_{t\in I^*}(xt-f^*(t))=f(x)$$

对于任意的 $x\in I$ 成立. 于是凸函数的勒让德变换是对合变换 (它的平方是恒等变换).

c) 成立不等式

$$xt\leqslant f(x)+f^*(t),\text{当}x\in I\text{、}t\in I^*\text{时}.$$

d) 当 f 是凸的可微函数时, $f^*(t)=tx_t-f(x_t)$, 其中 x_t 由方程

$$t=f'(x)$$

确定; 由此得到勒让德变换f^* 及其自变量 t 的几何解释, 它表明勒让德变换是一个在函数 f 图像的切线集合上定义的函数.

e) 当 $\alpha>1$ 且 $x\geqslant0$ 时, 函数 $f(x)=\frac{1}{\alpha}x^\alpha$ 的勒让德变换是函数 $f^*(t)=\frac{1}{\beta}t^\beta$, 其中 $t\geqslant0$ 且 $\frac{1}{\alpha}+\frac{1}{\beta}=1$; 由此, 注意到 c) 款, 得到已经知道的杨格不等式

$$xt\leqslant\frac{1}{\alpha}x^\alpha+\frac{1}{\beta}t^\beta.$$

---

①勒让德 (Legendre)(1752—1833) —— 著名的法国数学家.

f) 函数 $f(x)=e^x$ 的勒让德变换是函数 $f^*(t)=t\ln\dfrac{t}{e}, t>0$, 并且成立不等式

$$xt \leqslant e^x + t\ln\frac{t}{e}, \text{当} x\in\mathbb{R}, t>0\text{时}.$$

9. *曲线在一点处的曲率、曲率半径和曲率中心.* 设有一点在平面上运动. 其运动规律由一对关于时间二次可微的坐标函数 $x=x(t)$、$y=y(t)$ 确定. 这时它描画出一条曲线, 对此, 我们说这条曲线是以参数形式 $x=x(t)$、$y=y(t)$ 给定的. 由函数 $y=f(x)$ 的图像给出曲线的情形只是一种特殊参数形式, 此时, 可以认为 $x=t$ 而 $y=f(t)$. 我们想指出一个刻画曲线在一点处的弯曲程度的数, 就像圆半径的倒数可以作为圆周弯曲程度的指标. 我们下边使用的正是这个对比.

a) 求出点的加速度向量 $a(t)=(\ddot{x}(t),\ddot{y}(t))$ 的切向分量 a_t 和法向分量 a_n, 即把 a 表成和式 a_t+a_n, 其中向量 a_t 与速度向量 $v(t)=(\dot{x}(t),\dot{y}(t))$ 共线, 也就是说, 它的方向是沿轨迹切线的, 而向量 a_n 的方向是沿轨迹的法向的.

b) 证明: 当沿着半径为 r 的圆周运动时成立关系式

$$r=\frac{|v(t)|}{|a_n(t)|}.$$

c) 当沿着任意的曲线运动时, 考虑到 b), 自然把量

$$r(t)=\frac{|v(t)|}{|a_n(t)|}$$

叫做曲线在点 $(x(t),y(t))$ 处的*曲率半径*.

证明: 曲率半径按公式

$$r(t)=\frac{(\dot{x}^2+\dot{y}^2)^{3/2}}{|\dot{x}\ddot{y}-\dot{y}\ddot{x}|}$$

来计算.

d) 平面曲线在给定点 $(x(t)、y(t))$ 处的曲率半径的倒数叫做*绝对曲率*. 与绝对曲率一道, 还考虑量

$$k(t)=\frac{\dot{x}\ddot{y}-\ddot{x}\dot{y}}{(\dot{x}^2+\dot{y}^2)^{3/2}},$$

这个量叫做*曲率*.

证明: 曲率的符号表示曲线相对切线旋转的方向. 分析曲率有怎样的量纲.

e) 证明: 函数 $y=f(x)$ 的图像在点 $(x,f(x))$ 处的曲率可按公式

$$k(x)=\frac{y''(x)}{[1+(y')^2(x)]^{3/2}}$$

来计算. 研究 $k(x)$ 的符号, 亦即 $y''(x)$ 的符号与图像凸的方向的对应关系.

f) 选择常数 a、b、R 使得圆周 $(x-a)^2+(y-b)^2=R^2$ 与一条确定的以参数形式给出的曲线 $x=x(t), y=y(t)$ 在点 $x_0=x(t_0)$、$y_0=y(t_0)$ 处以尽可能高的阶数相切. 假定 $x(t)$、$y(t)$ 二次可微且 $(\dot{x}(t_0),\dot{y}(t_0))\neq(0,0)$.

上述圆周叫做曲线在点 (x_0,y_0) 处的*密切圆周*. 它的中心叫做曲线在点 (x_0,y_0) 处的*曲率中心*. 验证它的半径与在 c) 中定义的在这点处曲线的曲率半径相同.

g) 静止的质点在重力作用下从具抛物形剖面的冰山顶上开始下滑. 剖面方程是 $x+y^2=1, x\geqslant 0, y\geqslant 0$. 计算质点落地前的运动轨道.

§5. 复数 初等函数彼此间的联系

1. 复数 与在有理数域 $\mathbb{Q}$ 中代数方程 $x^2=2$ 没有解类似, 在实数域 $\mathbb{R}$ 中方程 $x^2=-1$ 也没有解. 同样类似的是, 我们引入了 $\mathbb{Q}$ 范围以外的符号 $\sqrt{2}$ 作为方程 $x^2=2$ 的解, 使它适合于 $\mathbb{Q}$ 中的运算, 并得到形如 $r_1+\sqrt{2}r_2$ 的新数, 其中 r_1、$r_2\in\mathbb{Q}$. 与此类似, 也可以引入符号 i 作为方程 $x^2=-1$ 的解, 并把这个 $\mathbb{R}$ 范围之外的数 i 和实数以及 $\mathbb{R}$ 中的算术运算联系起来.

实数域 $\mathbb{R}$ 的上述扩充有很多好处, 其中最大的优点是, 在所得的复数域 $\mathbb{C}$ 中, 任何实系数或复系数的代数方程再也不会没有解了.

现在按计划介绍本节内容.

a. 域 $\mathbb{R}$ 的代数扩充. 于是我们按欧拉的记号引入新的数 i ——虚单位, 使得 $i^2=-1$.

数 i 与实数间的相互作用的内容应是: 可以用数 $y\in\mathbb{R}$ 来乘 i, 即必须出现形如 iy 的数, 并且可以把这样的数与实数相加, 即出现形如 $x+iy$ 的数, 其中 x、$y\in\mathbb{R}$.

遵循高斯, 我们把形如 $x+iy$ 的对象叫做*复数*. 如果我们想在复数集上定义通常的满足交换律的加法运算和满足交换律并关于加法可分配的乘法运算, 那么作为定义必须令

$$(x_1+iy_1)+(x_2+iy_2):=(x_1+x_2)+i(y_1+y_2), \tag{1}$$

$$(x_1+iy_1)\cdot(x_2+iy_2):=(x_1x_2-y_1y_2)+i(x_1y_2+x_2y_1). \tag{2}$$

两个复数 x_1+iy_1 和 x_2+iy_2 当且仅当 $x_1=x_2$ 且 $y_1=y_2$ 时, 认为是相等的.

数 $x\in\mathbb{R}$ 与形如 $x+i0$ 的数恒等, 而 i 与数 $0+i\cdot1$ 恒等. 如同从 (1) 所见到的, 数 $0+i\cdot0=0\in\mathbb{R}$ 在复数集中起着零的作用, 如同从 (2) 所见到的, 数 $1+i\cdot0=1\in\mathbb{R}$ 起着单位的作用.

从实数的性质和定义 (1)、(2) 推出, 复数集是一个域, 它含有 $\mathbb{R}$ 为真子域.

复数域用符号 $\mathbb{C}$ 标记, 而它的元素通常用字母 z 和 w 表示.

对于 $\mathbb{C}$ 是域这一结论, 唯一不太明白, 从而需要验证的是, 任何异于零的复数 $z=x+iy$ 关于乘法都有逆 z^{-1}, 即 $z\cdot z^{-1}=1$. 我们来验证此事.

数 $x-iy$ 叫做数 $z=x+iy$ 的*共轭数*并以符号 $\bar{z}$ 记之.

我们注意到, 如果 $z\neq0$, 那么 $z\cdot\bar{z}=x^2+y^2+i\cdot0=x^2+y^2\neq0$. 因此, 应该取 $\dfrac{1}{x^2+y^2}\cdot\bar{z}=\dfrac{x}{x^2+y^2}-i\dfrac{y}{x^2+y^2}$ 作为 z^{-1}.

b. 域 $\mathbb{C}$ 的几何解释. 我们指出, 引入了复数的代数运算 (1)、(2) 之后, 那个把我们引向这些定义的符号 i 就不再是必须的了.

我们可以把复数 $z = x + iy$ 与有序实数对 (x, y) 等同起来, x、y 分别叫做复数 z 的实部和虚部 (记作 $x = \mathrm{Re}z, y = \mathrm{Im}z$[①]).

然而, 一旦把数对 (x, y) 当作平面 $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$ 的点的笛卡儿坐标, 就可以把复数与这个平面的点或坐标为 (x, y) 的二维向量等同起来.

在这种向量解释中, 复数的依坐标的加法 (1) 对应于向量的加法法则. 此外, 这样的解释势必引导出复数 z 的模 $|z|$ 的概念, 即把对应的向量 (x, y) 的模或长度作为复数 z 的模, 即

$$|z| = \sqrt{x^2 + y^2}, \text{若} z = x + iy. \tag{3}$$

同样地, 这种解释也引导出以复数 z_1、z_2 对应的平面上两点的距离作为这两个数之间的距离来量度的方法, 即以量

$$|z_1 - z_2| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{4}$$

作为复数 z_1、z_2 之间的距离这样一个量度方法.

复数的集合, 解释作平面的点的集合, 叫做*复平面*, 仍用符号 $\mathbb{C}$ 表示. 与此类似, 实数的集合与数轴同以符号 $\mathbb{R}$ 表示.

因为平面的点还可以通过极坐标 (r, φ) 给出, 极坐标 (r, φ) 与笛卡儿坐标经变换公式

$$\begin{aligned} x &= r\ \cos\varphi, \\ y &= r\ \sin\varphi, \end{aligned} \tag{5}$$

相联系, 所以复数

$$z = x + iy \tag{6}$$

也可以表示成

$$z = r(\cos\varphi + i\ \sin\varphi). \tag{7}$$

写法 (6) 和 (7) 分别叫做复数的*代数形式*和*三角形式*.

在写法 (7) 中, 数 $r \geqslant 0$ 叫做复数 z 的模 (因为, 从 (5) 可见, $r = |z|$), 而 φ 叫做数 z 的*辐角*. 辐角只当 $z \neq 0$ 时有意义. 根据函数 $\cos\varphi$ 和 $\sin\varphi$ 的周期性, 复数辐角的值彼此相差 2π 的倍数, 而符号 $\mathrm{Arg}\ z$ 表示形如 $\varphi + 2\pi k (k \in \mathbb{Z})$ 的角度的集合, 其中 φ 是某个满足关系式 (7) 的角度. 要想使复数单值地确定某个辐角 $\varphi \in \mathrm{Arg}\ z$, 那就要预先说定辐角取值的范围. 最常见的是取半开区间 $0 \leqslant \varphi < 2\pi$ 或 $-\pi < \varphi \leqslant \pi$ 来做这个限制范围. 如果做好了这样的选择, 就说选定了辐角的*分支* (或*主支*). 在选定的范围内, 辐角的值通常用 $\arg z$ 来表示.

---

① 来源于拉丁文 realis (实的) 和 imaginarius (虚的).

复数写法的三角公式 (7) 适合于做复数的乘法运算. 事实上, 若

$$z_1 = r_1(\cos\varphi_1 + i\ \sin\varphi_1),$$
$$z_2 = r_2(\cos\varphi_2 + i\ \sin\varphi_2),$$

则

$$\begin{aligned} z_1 \cdot z_2 &= (r_1\cos\varphi_1 + ir_1\sin\varphi_1)(r_2\cos\varphi_2 + ir_2\sin\varphi_2) \\ &= (r_1r_2\cos\varphi_1\cos\varphi_2 - r_1r_2\sin\varphi_1\sin\varphi_2) + \\ &\quad i(r_1r_2\sin\varphi_1\cos\varphi_2 + r_1r_2\cos\varphi_1\sin\varphi_2), \end{aligned}$$

即

$$z_1 \cdot z_2 = r_1r_2(\cos(\varphi_1 + \varphi_2) + i\sin(\varphi_1 + \varphi_2)). \tag{8}$$

于是, 当复数相乘时, 它们的模相乘而辐角相加.

我们指出, 我们实际上证明了, 若 $\varphi_1 \in \mathrm{Arg}\ z_1$ 且 $\varphi_2 \in \mathrm{Arg}\ z_2$, 则 $(\varphi_1 + \varphi_2) \in \mathrm{Arg}(z_1 \cdot z_2)$. 但由于辐角之确定精确到 $2\pi k$, 所以有

$$\mathrm{Arg}\ z_1 \cdot z_2 = \mathrm{Arg}\ z_1 + \mathrm{Arg}\ z_2, \tag{9}$$

这是一个集合等式, 右边是形如 $\varphi_1 + \varphi_2$ 的数的全体, 其中 $\varphi_1 \in \mathrm{Arg}\ z_1$、$\varphi_2 \in \mathrm{Arg}\ z_2$. 因此, 按等式 (9) 的意义来理解辐角的和是有效的.

例如, 当这样来理解辐角的等式的时候, 可以作出这样的断语: 两个复数相等, 当且仅当它们的模和辐角分别相等.

从公式 (8) 按归纳法, 显然可以推出下列棣莫弗①公式:

若 $z = r(\cos\varphi + i\sin\varphi)$, 则

$$z^n = r^n(\cos n\varphi + i\sin n\varphi). \tag{10}$$

考虑到关于复数辐角的解释, 可以用棣莫弗公式明确写出方程 $z^n = a$ 的全部复数解.

实际上, 如果

$$a = \rho(\cos\psi + i\sin\psi),$$

且据公式 (10)

$$z^n = r^n(\cos n\varphi + i\sin n\varphi),$$

那么 $r = \sqrt[n]{\rho}$ 且 $n\varphi = \psi + 2k\pi, k \in \mathbb{Z}$, 由此 $\varphi_k = \dfrac{\psi}{n} + \dfrac{2\pi}{n}k$. 显然, 仅当 $k = 0, 1, \cdots, n-1$ 时得到不同的复数. 于是, 我们求得复数 a 的 n 个不同的根:

$$z_k = \sqrt[n]{\rho}\left(\cos\left(\frac{\psi}{n} + \frac{2\pi}{n}k\right) + i\sin\left(\frac{\psi}{n} + \frac{2\pi}{n}k\right)\right), k = 0, 1, \cdots, n-1.$$

①棣莫弗 (Abraham de Moivre) (1667—1754) —— 英国数学家, 生于法国.

特别地, 若 $a=1$ 即 $\rho=1$ 且 $\psi=0$, 则有

$$z_k=\sqrt[n]{1}=\cos\left(\frac{2\pi}{n}k\right)+i\sin\left(\frac{2\pi}{n}k\right)\quad (k=0,1,\cdots,n-1).$$

这些点位于单位圆周上, 构成正 n 角形的顶点.

与复数本身的几何解释相关, 提一下复数之间的算术运算的几何解释是有好处的.

对于固定的 $b\in\mathbb{C}$, 和式 $z+b$ 可以解释作由公式 $z\mapsto z+b$ 给出的 $\mathbb{C}$ 到自身的映射. 它是在平面上平移向量 b 的运算.

对于固定的 $a=|a|(\cos\varphi+i\sin\varphi)\neq 0$, 积 $a\cdot z$ 可以解释为 $\mathbb{C}$ 到自身的映射 $z\mapsto a\cdot z$. 它是伸长到 $|a|$ 倍并旋转一个角度 $\varphi\in \operatorname{Arg} a$ 的复合. 这可从公式 (8) 看出来.

2. C 中的收敛及复数项级数 复数之间的距离 (4) 使我们可以把数 $z_0\in\mathbb{C}$ 的 ε 邻域定义为集合

$$\{z\in\mathbb{C}\,|\,|z-z_0|<\varepsilon\}.$$

如果 $z_0=x_0+iy_0$ 的话, 这个集合就是一个以点 (x_0,y_0) 为中心以 ε 为半径的 (不包括圆周的) 圆.

我们说复数序列 $\{z_n\}$ 收敛到数 $z_0\in\mathbb{C}$ 指的是

$$\lim_{n\to\infty}|z_n-z_0|=0.$$

从不等式

$$\max\{|x_n-x_0|,|y_n-y_0|\}\leqslant|z_n-z_0|\leqslant|x_n-x_0|+|y_n-y_0| \tag{11}$$

看出, 复数序列收敛当且仅当它的项的实部的序列和虚部的序列都收敛.

与实数列的情形类似, 复数列 $\{z_n\}$ 叫做*基本列*或*柯西列*, 如果对于任意的 $\varepsilon>0$, 都存在号码 $N\in\mathbb{N}$, 使当 m、$n>N$ 时, $|z_m-z_n|<\varepsilon$ 成立.

从不等式 (11) 看到, 复数列是基本的当且仅当所给数列的项的实部列和虚部列都是基本的.

因此, 考虑到实数列的柯西准则, 我们根据 (11) 断定成立如下

命题 1 (柯西准则) *复数列收敛当且仅当它是基本的.*

如果把复数的级数

$$z_1+z_2+\cdots+z_n+\cdots \tag{12}$$

的和理解为它的部分和 $S_n=z_1+\cdots+z_n$ 当 $n\to\infty$ 时的极限的话, 那么我们也得到复数级数收敛的柯西准则.

命题 2 级数 (12) 收敛当且仅当对于任意的 $\varepsilon > 0$, 存在数 $N \in \mathbb{N}$, 使得对于任意的 $n > m > N$ 有

$$|z_m + z_{m+1} + \cdots + z_n| < \varepsilon. \tag{13}$$

由此看出, 要使级数收敛, 必须当 $n \to \infty$ 时 $z_n \to 0$. 其实从级数 (12) 收敛的定义本身就已看出此事.

与实的情形一样, 级数 (12) 叫做绝对收敛的, 如果级数

$$|z_1| + |z_2| + \cdots + |z_n| + \cdots \tag{14}$$

收敛的话.

从柯西准则和不等式

$$|z_m + z_{m+1} + \cdots + z_n| \leqslant |z_m| + |z_{m+1}| + \cdots + |z_n| \quad (m < n)$$

推出, 若级数 (12) 绝对收敛, 则它收敛.

例 级数

1) $1 + \frac{1}{1!}z + \frac{1}{2!}z^2 + \cdots + \frac{1}{n!}z^n + \cdots$,

2) $z - \frac{1}{3!}z^3 + \frac{1}{5!}z^5 - \cdots$,

3) $1 - \frac{1}{2!}z^2 + \frac{1}{4!}z^4 - \cdots$

对于任意的 $z \in \mathbb{C}$ 都绝对收敛, 因为我们已经知道级数

1′) $1 + \frac{1}{1!}|z| + \frac{1}{2!}|z|^2 + \cdots$,

2′) $|z| + \frac{1}{3!}|z|^3 + \frac{1}{5!}|z|^5 + \cdots$,

3′) $1 + \frac{1}{2!}|z|^2 + \frac{1}{4!}|z|^4 + \cdots$

对于任何值 $|z| \in \mathbb{R}$ 都收敛. 我们指出, 此处使用了等式 $|z|^n = |z^n|$.

例 4 级数 $1 + z + z^2 + \cdots$ 当 $|z| < 1$ 时绝对收敛, 且它的和等于 $s = \frac{1}{1-z}$. 当 $|z| \geqslant 1$ 时它不收敛, 因为此时级数的一般项不趋于零.

形如

$$c_0 + c_1(z - z_0) + \cdots + c_n(z - z_0)^n + \cdots \tag{15}$$

的级数叫做*幂级数*.

用柯西判别法 (见第三章 §1 第 4 段) 于级数

$$|c_0| + |c_1(z - z_0)| + \cdots + |c_n(z - z_0)^n| + \cdots, \tag{16}$$

我们断定, 当

$$|z-z_0|<\left(\overline{\lim_{n\to\infty}}\sqrt[n]{|c_n|}\right)^{-1}$$

时, 级数收敛, 当 $|z-z_0|>\left(\overline{\lim\limits_{n\to\infty}}\sqrt[n]{|c_n|}\right)^{-1}$ 时它的一般项不趋于零.

由此得到如下

命题 3 (柯西-阿达马[①]公式) *幂级数(15) 在以点 z_0 为中心以 R 为半径的圆 $|z-z_0|<R$ 内收敛, 其中 R (叫收敛半径) 按下列柯西-阿达马公式确定:*

$$R=\left(\overline{\lim_{n\to\infty}}\sqrt[n]{|c_n|}\right)^{-1}. \tag{17}$$

在这个圆的外部的任何点处幂级数都发散.

在这个圆的任何点 (内点) 处幂级数都绝对收敛.

注 关于在周界 $|z-z_0|=R$ 上的收敛性, 命题 3 一无所述, 因为此处一切逻辑上容许的情况都可能发生.

例 级数

5) $\sum\limits_{n=1}^{\infty}z^n$, 6) $\sum\limits_{n=1}^{\infty}\frac{1}{n}z^n$, 7) $\sum\limits_{n=1}^{\infty}\frac{1}{n^2}z^n$

在单位圆 $|z|<1$ 中收敛. 但级数 5) 在 $|z|=1$ 上处处发散; 级数 6) 当 $z=1$ 时发散而当 $z=-1$ 时可以证明是收敛的; 级数 7) 当 $|z|=1$ 时绝对收敛, 因为 $\left|\frac{1}{n^2}z^n\right|=\frac{1}{n^2}$.

应该注意到在命题 3 的叙述中未加考虑但可能蜕化的情形, 即在公式 (17) 中 $R=0$ 的情形. 此时, *整个收敛圆*蜕化成级数 (15) 的唯一的*收敛点* z_0.

从命题 3 显然推出

推论 (关于幂级数的阿贝尔第一定理) *若幂级数 (15) 对于某个值 z^* 收敛, 则它对于任意满足 $|z-z_0|<|z^*-z_0|$ 的 z 都绝对收敛.*

目前得到的这些命题都可看作是我们已经知道的事实的简单推广. 现在我们来证明两个关于级数的一般的命题. 这两个命题我们以前不曾以任何形式论证过, 虽然曾部分地讨论过它们所涉及的问题.

命题 4 *若复数级数 $z_1+z_2+\cdots+z_n+\cdots$ 绝对收敛, 则重排它的项所得的级数 $z_{m_1}+z_{m_2}+\cdots+z_{m_k}+\cdots$ [②]同样绝对收敛, 且收敛到同一个和.*

---

①阿达马 (Hadamard) (1865—1963) —— 著名的法国数学家.

②第二个级数号码为 k 的项 (第 k 项) 是原来级数的号码为 m_k 的项. 映射 $\mathbb{N}\ni k\mapsto m_k\in\mathbb{N}$ 假定是自然数集 $\mathbb{N}$ 到自身的满单射.

◀ 考虑到级数 $\sum_{n=1}^{\infty}|z_n|$ 的收敛性, 对于数 $\varepsilon > 0$, 我们找到号码 $N \in \mathbb{N}$, 使得 $\sum_{n=N+1}^{\infty}|z_n| < \varepsilon$.

另外, 找得到号码 $K \in \mathbb{N}$, 使当 $k > K$ 时, 在和式 $\tilde{s}_k = a_{n_1} + \cdots + a_{n_k}$ 的加数中包含着和式 $s_N = a_1 + \cdots + a_N$ 的所有的加数. 如果 $s = \sum_{n=1}^{\infty} z_n$, 那么, 我们得到当 $k > K$ 时

$$|s - \tilde{s}_k| \leqslant |s - s_N| + |s_N - \tilde{s}_k| \leqslant \sum_{n=N+1}^{\infty}|z_n| + \sum_{n=N+1}^{\infty}|z_n| < 2\varepsilon.$$

这样就证明了当 $k \to \infty$ 时, $\tilde{s}_k \to s$. 若应用已证之事于级数 $|z_1|+|z_2|+\cdots+|z_n|+\cdots$ 及级数 $|z_{m_1}| + |z_{m_2}| + \cdots + |z_{m_k}| + \cdots$, 则得知后一级数收敛. 这就完全证明了命题 4. ▶

下一个命题是关于级数的乘积

$$(a_1 + a_2 + \cdots + a_n + \cdots)(b_1 + b_2 + \cdots + b_n + \cdots)$$

的. 问题在于, 如果我们要打开括弧并写出一切可能的两两乘积 a_ib_j, 那么, 对于这些乘积并不存在一个自然的求和次序, 因为有两个求和指标. 我们知道, 对于 i、$j \in \mathbb{N}$, 数对 (i, j) 的集合是可数的, 因此可以按某种次序写出以 a_ib_j 为项的级数. 此级数的和可能与这些项在级数中的排列次序有关. 但是刚才我们已经见到, 在绝对收敛的情形, 级数的和不依赖于项的排列次序. 所以, 最好还是弄清楚以 a_ib_j 为项的级数什么时候绝对收敛.

命题 5 绝对收敛的级数的乘积是绝对收敛的级数, 它的和等于作为乘数的级数的和的乘积.

◀ 先指出, 不管我们取怎样一个其项形如 a_ib_j 的有限和 $\sum a_ib_j$, 总可以指出这样的 N, 使得和

$$A_N = a_1 + \cdots + a_N \text{与和} B_N = b_1 + \cdots + b_N$$

的乘积包含原来所取的和的全部被加项. 因此

$$\left|\sum a_ib_j\right| \leqslant \sum |a_ib_j| \leqslant \sum_{i,j=1}^{N}|a_ib_j| = \sum_{i=1}^{N}|a_i| \cdot \sum_{j=1}^{N}|b_j| \leqslant \sum_{i=1}^{\infty}|a_i| \cdot \sum_{j=1}^{\infty}|b_j|.$$

由此推出级数 $\sum_{i,j=1}^{\infty} a_ib_j$ 的绝对收敛性, 从而它的和是单值确定的, 与被加项的次序无关. 例如这时可以用和 $A_n = a_1 + \cdots + a_n$, $B_n = b_1 + \cdots + b_n$ 的乘积当 $n \to \infty$ 时

的极限来得到它. 但当 $n \to \infty$ 时, $A_n \cdot B_n \to AB$, 其中 $A = \sum_{n=1}^{\infty} a_n, B = \sum_{n=1}^{\infty} b_n$, 这样就完成了所述命题 5 的证明. ▶

举一个重要的例子.

例 8　级数 $\sum_{n=0}^{\infty} \frac{1}{n!}a^n, \sum_{m=0}^{\infty} \frac{1}{m!}b^m$ 都绝对收敛. 在计算这两个级数的积时, 我们把幂指数的和 $m+n$ 等于 k 的项 $a^n \cdot b^m$ 并在一起, 就得到级数

$$\sum_{k=0}^{\infty}\left(\sum_{n+m=k} \frac{1}{n!}a^n \cdot \frac{1}{m!}b^m\right).$$

然而

$$\sum_{m+n=k} \frac{1}{n!m!}a^n \cdot b^m = \frac{1}{k!}\sum_{n=0}^{k} \frac{k!}{n!(k-n)!}a^n b^{k-n} = \frac{1}{k!}(a+b)^k,$$

所以我们得到

$$\sum_{n=0}^{\infty} \frac{1}{n!}a^n \cdot \sum_{m=0}^{\infty} \frac{1}{m!}b^m = \sum_{k=0}^{\infty} \frac{1}{k!}(a+b)^k. \tag{18}$$

3. 欧拉[①]公式以及初等函数彼此间的联系　在例 1)—3) 中我们确定了三个级数在复数域 $\mathbb{C}$ 中的绝对收敛性. 这三个级数分别是由定义在 $\mathbb{R}$ 上的函数 e^x、$\sin x$、$\cos x$ 的泰勒展开式推广到复数域得到的. 由于这个缘故, 函数 e^z、$\cos z$、$\sin z$ 在 $\mathbb{C}$ 中的下述定义是很自然的:

$$e^z = \exp z := 1 + \frac{1}{1!}z + \frac{1}{2!}z^2 + \frac{1}{3!}z^3 + \cdots, \tag{19}$$

$$\cos z := 1 - \frac{1}{2!}z^2 + \frac{1}{4!}z^4 - \cdots, \tag{20}$$

$$\sin z := \frac{1}{1!}z - \frac{1}{3!}z^3 + \frac{1}{5!}z^5 - \cdots. \tag{21}$$

遵循欧拉的办法, 在 (19) 中代入 $z = iy$, 以适当的方式把此时所得级数的部分和的被加项分组, 求得

$$1 + \frac{1}{1!}(iy) + \frac{1}{2!}(iy)^2 + \frac{1}{3!}(iy)^3 + \frac{1}{4!}(iy)^4 + \frac{1}{5!}(iy)^5 + \cdots$$
$$= \left(1 - \frac{1}{2!}y^2 + \frac{1}{4!}y^4 - \cdots\right) + i\left(\frac{1}{1!}y - \frac{1}{3!}y^3 + \frac{1}{5!}y^5 - \cdots\right),$$

即

$$\boxed{e^{iy} = \cos y + i\sin y.} \tag{22}$$

---

①欧拉 (Euler) (1707—1783)——卓越的数学家和力学家, 按族系说是瑞士人, 一生大部分时间住在彼得堡. 按拉普拉斯的说法,"欧拉是 18 世纪后半叶全体数学家共同的导师".

这就是著名的欧拉公式.

在推导中我们用到了 $i^2=-1$、$i^3=-i$、$i^4=1$、$i^5=i$ 等等. 在公式 (22) 中, 数 y 既可以是实数也可以是任意的复数.

从定义 (20)、(21) 见到

$$\cos(-z)=\cos z,$$
$$\sin(-z)=-\sin z,$$

即 $\cos z$ 是偶函数, $\sin z$ 是奇函数. 于是

$$e^{-iy}=\cos y-i\sin y.$$

将此式与公式 (22) 对比, 得到

$$\cos y=\frac{1}{2}(e^{iy}+e^{-iy}),$$
$$\sin y=\frac{1}{2i}(e^{iy}-e^{-iy}).$$

由于 y 是任意复数, 所以这些等式最好还是改用不引起误解的记号写出:

$$\begin{aligned}\cos z&=\frac{1}{2}(e^{iz}+e^{-iz}),\\ \sin z&=\frac{1}{2i}(e^{iz}-e^{-iz}).\end{aligned}\tag{23}$$

这样一来, 如果设 $\exp z$ 由关系式 (19) 确定, 则公式 (23) (等价于展开式 (20)、(21)) 以及公式

$$\begin{aligned}\operatorname{ch} z&=\frac{1}{2}(e^{z}+e^{-z}),\\ \operatorname{sh} z&=\frac{1}{2}(e^{z}-e^{-z}),\end{aligned}\tag{24}$$

相应地可以取作圆函数和双曲函数的定义. 假若我们把那些关于三角函数 的有启发性的然而有时不十分严格有据的看法统统忘记 (尽管这些看法把我们引导到欧拉公式), 现在可以运用典型的数学技巧: 取公式 (23)、(24) 作为定义, 从这些定义可以完全从逻辑上得出圆函数和双曲函数的性质.

例如, 基本恒等式

$$\cos^2 z+\sin^2 z=1,$$
$$\operatorname{ch}^2 z-\operatorname{sh}^2 z=1,$$

以及偶性, 都可直接验证.

更深刻的性质, 如像和的余弦及正弦的公式, 可以从指数函数的特征性质

$$\exp(z_1+z_2)=\exp z_1\cdot\exp z_2\tag{25}$$

推出, 性质 (25) 显然可从定义 (19) 及公式 (18) 得出. 我们来推导和的余弦及正弦的公式.

一方面按欧拉公式

$$e^{i(z_1+z_2)} = \cos(z_1+z_2) + i\sin(z_1+z_2); \tag{26}$$

另一方面, 按指数函数的性质及欧拉公式有

$$\begin{aligned} e^{i(z_1+z_2)} &= e^{iz_1} \cdot e^{iz_2} = (\cos z_1 + i\sin z_1)(\cos z_2 + i\sin z_2) \\ &= (\cos z_1 \cos z_2 - \sin z_1 \cdot \sin z_2) + i(\sin z_1 \cos z_2 + \cos z_1 \sin z_2). \end{aligned} \tag{27}$$

如果 z_1、z_2 都是实数, 那么, 使公式 (26) 和 (27) 中的数的实部、虚部分别相等, 我们就得到所求的公式. 由于我们打算对任意的 z_1、$z_2 \in \mathbb{C}$ 来证明这些公式, 所以使用 $\cos z$ 的偶性和 $\sin z$ 的奇性, 再写下一个等式:

$$e^{-i(z_1+z_2)} = (\cos z_1 \cos z_2 - \sin z_1 \sin z_2) - i(\sin z_1 \cos z_2 + \cos z_1 \sin z_2). \tag{28}$$

比较 (27) 和 (28), 求得

$$\begin{aligned} \cos(z_1+z_2) &= \frac{1}{2}(e^{i(z_1+z_2)} + e^{-i(z_1+z_2)}) \\ &= \cos z_1 \cdot \cos z_2 - \sin z_1 \sin z_2, \\ \sin(z_1+z_2) &= \frac{1}{2i}(e^{i(z_1+z_2)} - e^{-i(z_1+z_2)}) \\ &= \sin z_1 \cdot \cos z_2 + \cos z_1 \sin z_2. \end{aligned}$$

完全类似地可以得到关于双曲函数 $\operatorname{ch} z$ 和 $\operatorname{sh} z$ 的相应的公式. 顺便说一下, 从公式 (23)、(24) 可见, 函数 $\operatorname{ch} z$、$\operatorname{sh} z$ 与函数 $\cos z$、$\sin z$ 通过简单的关系式

$$\operatorname{ch} z = \cos iz,$$

$$\operatorname{sh} z = -i \sin iz$$

相互联系着.

但是, 要从定义 (23) 得到像 $\sin\pi = 0$ 或 $\cos(z+2\pi) = \cos z$ 这样几何上十分明显的事实已是十分困难的了. 这就表示, 当致力于精确性的时候, 仍不应忘记相应的函数是在哪些问题中自然地产生出来的. 由于这个缘故, 在此我们不打算去克服那些在描述三角函数的性质时可能发生的与定义 (23) 有关的困难, 而是在积分理论之后再回过头来讨论这些函数. 此刻我们的目的只是要展示看来似乎完全不同的函数所具有的那种美妙的统一性, 这种统一性不进入复数域是发现不了的.

如果认为对于 $x \in \mathbb{R}$,

$$\cos(x+2\pi) = \cos x, \quad \sin(x+2\pi) = \sin x, \quad \cos 0 = 1, \quad \sin 0 = 0$$

都是已知的, 那么从欧拉公式 (22) 得到关系式

$$\boxed{e^{i\pi}+1=0,} \tag{29}$$

在这个式子里出现了各不同的数学领域里的最重要的常数 (1 —— 算术, π—— 几何, e —— 分析, i—— 代数).

从 (25) 和 (29), 像从公式 (22) 一样看出

$$\exp(z+i2\pi)=\exp z,$$

也就是说, 在 $\mathbb{C}$ 中指数函数原来是一个周期函数, 函数的周期 $T=i2\pi$, 是纯常数.

注意到欧拉公式, 现在可以把复数的三角写法 (7) 表示成

$$z=re^{i\varphi},$$

其中 r 是数 z 的模, 而 φ 是它的辐角.

棣莫弗公式现在变得十分简单:

$$z^n=r^ne^{in\varphi}. \tag{30}$$

4. 函数的幂级数表示, 解析性 定义在某个集 $E\subset\mathbb{C}$ 上的复变量 z 的复值函数$w=f(z)$ 是一个 E 到 $\mathbb{C}$ 的映射 $f:E\to\mathbb{C}$. 这种函数的图像是 $\mathbb{C}\times\mathbb{C}=\mathbb{R}^2\times\mathbb{R}^2=\mathbb{R}^4$ 中的子集合, 因此不具有通常的直观性. 为了弥补这种损失, 通常取两张复平面, 在一张上标示定义域的点, 而在另一张上标示值域的点.

在下面的几个例子中, 标出了定义域 E 和它在相应的映射下的像.

例 9

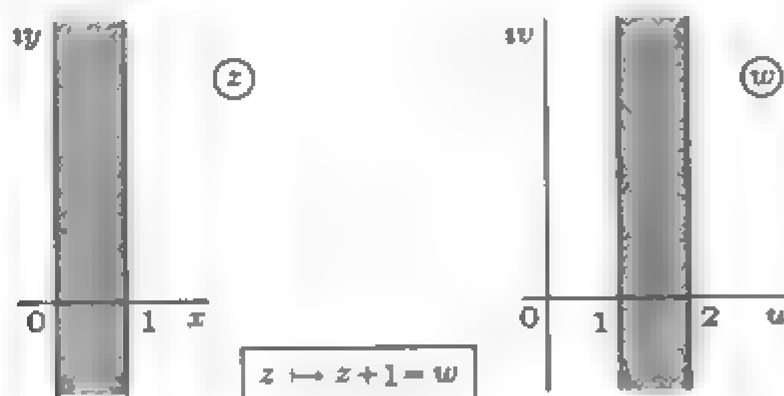


图 37

例 10

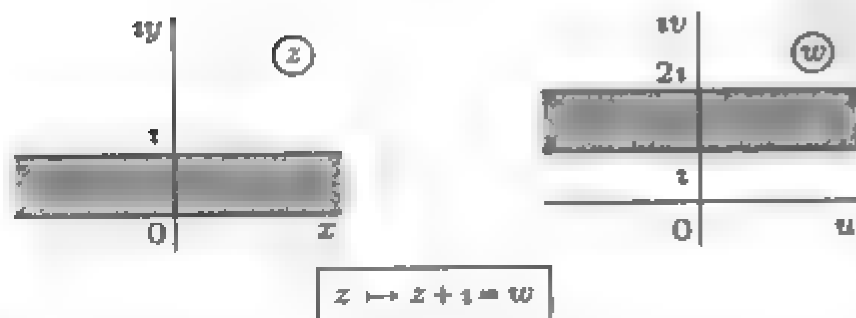


图 38

例 11

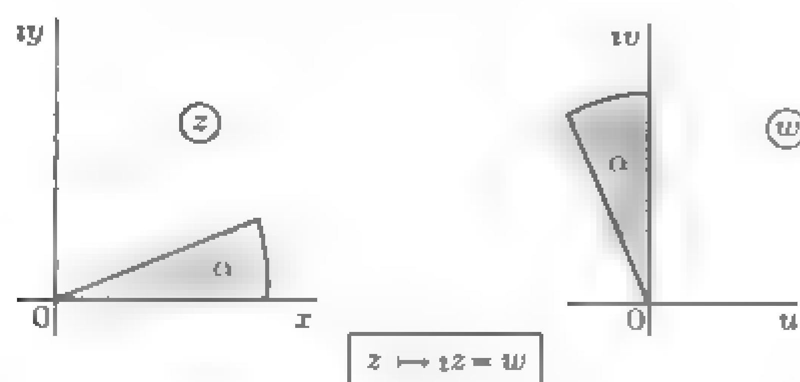


图 39

这从 $i = e^{i\frac{\pi}{2}}, z = re^{i\varphi}$ 从而 $iz = re^{i(\varphi+\frac{\pi}{2})}$ 推知, 这也就是说从旋转 $\frac{\pi}{2}$ 角得到.

例 12

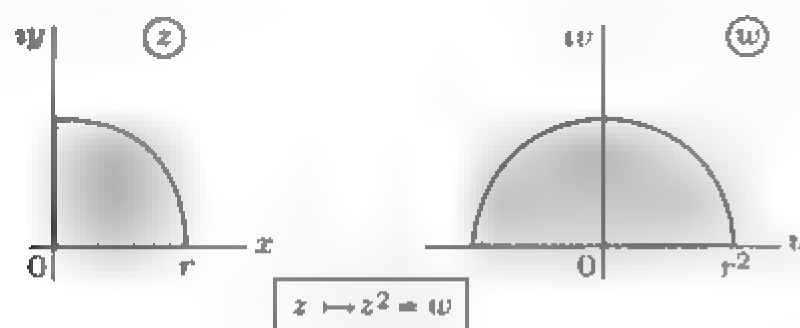


图 40

因为如果 $z = re^{i\varphi}$, 则 $z^2 = r^2e^{i2\varphi}$.

例 13

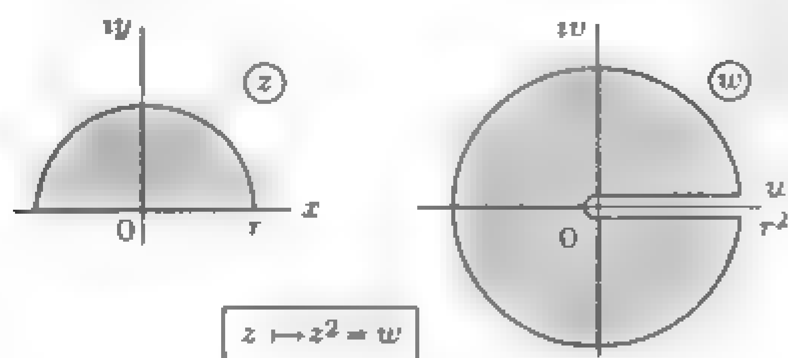


图 41

例 14

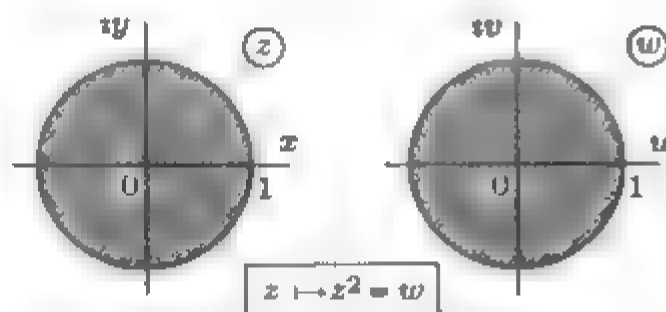


图 42

从例 13,14 明白, 在这种情况下, 单位圆的像仍是单位圆, 只是叠成两层.

例 15

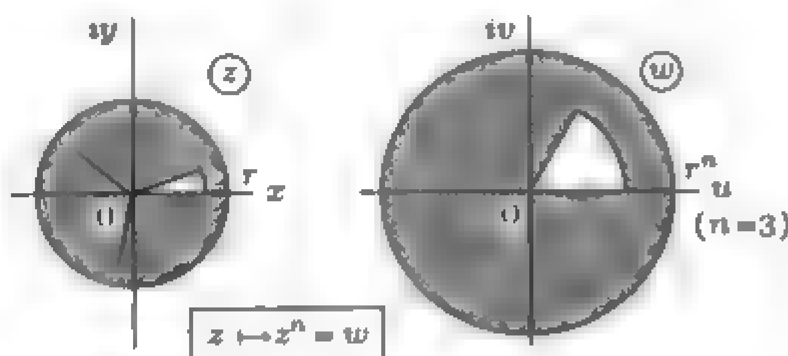


图 43

若 $z = re^{i\varphi}$, 则根据 (30)$z^n = r^n e^{in\varphi}$, 所以在我们的情况下, 半径为 r 的圆的像是半径为 r^n 的圆, 且后者的每个点都是原来的圆的 n 个点的像 (顺便说一下, 这 n 个点分布在一个正 n 角形的顶点处).

这里例外的只是点 $w = 0$, 它的原像是点 $z = 0$.

不过, 当 $z \to 0$ 时, 函数 z^n 是 n 阶无穷小量, 因此说 $z = 0$ 是函数的 n 阶零点. 考虑到零的这个重数, 现在可以说, 在映射 $z \mapsto z^n = w$ 之下, 任一点 w 的原像的数目都等于 n. 特别地, 方程 $z^n = 0$ 有 n 个重叠的根 $z_1 = z_2 = \cdots = z_n = 0$.

相应于连续性的一般定义, $f(z)$ 叫做在点 $z_0 \in \mathbb{C}$ 的连续复变函数, 如果对于它的值 $f(z_0)$ 的任何邻域 $V(f(z_0))$, 存在这样的邻域 $U(z_0)$, 使得对于任意的 $z \in U(z_0)$, 有

$$f(z) \in V(f(z_0)),$$

简单地说就是

$$\lim_{z \to z_0} f(z) = f(z_0).$$

像对于实的情形那样, 把量

$$f'(z_0) = \lim_{z \to z_0} \frac{f(z) - f(z_0)}{z - z_0} \tag{31}$$

叫做复变函数 $f(z)$ 在点 z_0 处的导数, 如果这个极限存在的话.

等式 (31) 等价于当 $z \to z_0$ 时

$$f(z) - f(z_0) = f'(z_0)(z - z_0) + o(z - z_0). \tag{32}$$

此等式相当于函数在点 z_0 处可微的定义.

由于可微复变函数的定义与实函数的相应定义是一致的, 而域 $\mathbb{C}$ 的算术性质也与 $\mathbb{R}$ 的一样, 所以, 一般的微分法则在复的情形下也都成立.

例 16

$$(f + g)'(z) = f'(z) + g'(z),$$
$$(f \cdot g)'(z) = f'(z)g(z) + f(z) \cdot g'(z),$$
$$(g \circ f)'(z) = g'(f(z)) \cdot f'(z).$$

因此, 若 $f(z) = z^2$, 则 $f'(z) = 1 \cdot z + z \cdot 1 = 2z$. 而若 $f(z) = z^n$, 则 $f'(z) = nz^{n-1}$. 又若

$$P_n(z) = c_0 + c_1(z - z_0) + \cdots + c_n(z - z_0)^n,$$

则

$$P_n'(z) = c_1 + 2c_2(z - z_0) + \cdots + nc_n(z - z_0)^{n-1}.$$

下述定理给出了幂级数的微分法.

定理 1 *幂级数 $f(z) = \sum\limits_{n=0}^{\infty} c_n(z - z_0)^n$ 的和是定义在其收敛圆内的无穷可微函数, 而且, 有*

$$f^{(k)}(z) = \sum_{n=0}^{\infty} \frac{d^k}{dz^k}(c_n(z - z_0)^n), k = 0, 1, \cdots$$

以及

$$c_n = \frac{f^{(n)}(z_0)}{n!}, n = 0, 1, \cdots$$

◀ 在 $f^{(k)}(z)$ 的表达式中令 $k = n$ 和 $z = z_0$ 立刻得出系数 c_n 的表达式.

而 $f^{(k)}(z)$ 的公式本身, 只要对 $k = 1$ 的情形加以验证即可, 因为 $f'(z)$ 也是幂级数和的形式.

下边验证, 函数 $\varphi(z) = \sum\limits_{n=1}^{\infty} nc_n(z - z_0)^{n-1}$ 实际就是 $f(z)$ 的导数.

首先, 我们注意到, 根据柯西 — 阿达马公式 (17), 最后这个级数的收敛半径与原来的 $f(z)$ 的幂级数的收敛半径相同.

为了书写简单起见, 下边将认为 $z_0 = 0$, 即 $f(z) = \sum\limits_{n=0}^{\infty} c_n z^n, \varphi(z) = \sum\limits_{n=1}^{\infty} nc_n z^{n-1}$, 而且, 当 $|z| < R$ 时级数收敛.

由于对收敛圆内幂级数绝对收敛. 可以发现, 最重要的是, 当 $|z| \leqslant r < R$ 时成立估计 $|nc_n z^{n-1}| = n|c_n||z^{n-1}| \leqslant n|c_n|r^{n-1}$, 而级数 $\sum\limits_{n=1}^{\infty} n|c_n|r^{n-1}$ 是收敛的. 因此, 对于任意的 $\varepsilon > 0$ 存在号码 N 使当 $|z| \leqslant r$ 时有

$$\left|\sum_{n=N+1}^{\infty} nc_n z^{n-1}\right| \leqslant \sum_{n=N+1}^{\infty} n|c_n|r^{n-1} \leqslant \frac{\varepsilon}{3}.$$

这样一来, 在圆 $|z| < r$ 中任何一点, 函数 $\varphi(z) < r$ 确定它的级数的前 N 项部分和之差不超过 $\frac{\varepsilon}{3}$.

现设 ς 和 z 是这个圆的任意两点. 与上边类似, 由变换

$$\frac{f(\varsigma) - f(z)}{\varsigma - z} = \sum_{n=1}^{\infty} c_n \frac{\varsigma^n - z^n}{\varsigma - z} = \sum_{n=1}^{\infty} c_n(\varsigma^{n-1} + \varsigma^{n-2}z + \cdots + \varsigma z^{n-2} + z^{n-1})$$

和估计 $|c_n(\varsigma^{n-1} + \cdots + z^{n-1})| \leqslant |c_n|nr^{n-1}$, 可得, 我们感兴趣的差比, 当 $|\varsigma| < r$ 和 $|z| < r$ 时, 与确定它的级数的前 N 项部分和之差不超过 $\varepsilon/3$. 因此, 当 $|\varsigma| < r$ 和 $|z| < r$ 时, 有

$$\left|\frac{f(\varsigma) - f(z)}{\varsigma - z} - \varphi(z)\right| \leqslant \left|\sum_{n=1}^{N} c_n \frac{\varsigma^n - z^n}{\varsigma - z} - \sum_{n=1}^{N} nc_n z^{n-1}\right| + \frac{2}{3}\varepsilon.$$

如果现在固定 z 并令 ς 趋于 z, 那么, 在有限和中取极限, 我们看到, 当 ς 与 z 充分接近时, 最后这个不等式右边部分, 从而其左边将小于 ε.

这样一来, 由 $r < R$ 的任意性, 证明了, 对于圆 $|z| < R$ 中任一点 z, 成立 $f'(z) = \varphi(z)$. ►

这个定理正好使得可以确定这样一个函数类, 它的每个函数的泰勒级数都收敛到这个函数本身.

说函数在点 $z_0 \in \mathbb{C}$ 处解析, 指的是它在这点的某个邻域内可以表示成下述 ("解析的") 形式:

$$f(z) = \sum_{n=0}^{\infty} c_n (z - z_0)^n,$$

即关于 $z - z_0$ 的幂级数的和.

不难验证 (参见问题 7), 幂级数的和在它的收敛圆的任意内点解析.

考虑到函数解析性的定义, 从已证过的定理得到:

推论　a) 函数如果在一点解析, 那么, 它在该点无穷次可微且其泰勒级数在该点的一个邻域中收敛于它.

b) 定义在一点邻域中且在该点无穷次可微的函数的泰勒级数. 在该点某邻域中收敛于这个函数, 当且仅当, 函数在该点解析.

在复变函数论中将证明一个极妙的事实, 即若函数 $f(z)$ 在点 z_0 的邻域内可微, 则它就在这点解析. 对于实函数并无类似之事. 这确实是令人吃惊的, 因为根据上边已证的定理, 由此推出, 如果函数 $f(z)$ 在一点的邻域内有一阶导数 $f'(z)$, 那么, 在这个邻域内它同样有任意阶的导数.

此事初看起来出乎意料, 有点像当我们把一个具体的方程 $z^2=-1$ 的根 i 并入 $\mathbb{R}$ 并得到域 $\mathbb{C}$, 在 $\mathbb{C}$ 中任何代数多项式 $P(z)$ 就都有根了. 代数方程 $P(z)=0$ 在 $\mathbb{C}$ 中可解这一事实, 我们打算用它, 对它进行证明, 作为说明本节引入的复数及复变函数初步概念的一个很好的实例.

5. 复数域 $\mathbb{C}$ 的代数封闭性　如果我们能证明任何复系数多项式

$$P(z)=c_0+c_1z+\cdots+c_nz^n, n\geqslant 1,$$

在 $\mathbb{C}$ 中都有根, 那就再也不会由于某个代数方程在 $\mathbb{C}$ 中不可解而去扩充域 $\mathbb{C}$. 从这个意义上说, 任何多项式 $P_n(z)$ 都有根这样一个命题确定了域 $\mathbb{C}$ 的代数封闭性.

为什么在 $\mathbb{C}$ 中一下子任何多项式都有根, 而在 $\mathbb{R}$ 中却可以没有根呢? 为了对此有一个完全直观的概念, 我们使用复数和复变函数的几何解释.

注意到

$$P(z)=z^n\left(\frac{c_0}{z^n}+\frac{c_1}{z^{n-1}}+\cdots+\frac{c_{n-1}}{z}+c_n\right).$$

可知, 当 $|z|\to\infty$ 时, $P(z)=c_nz^n+o(z^n)$. 由于使我们感兴趣的是方程 $P(z)=0$ 的根, 所以只要以 c_n 除方程的两边, 就可以认为多项式 $P(z)$ 的系数 $c_n=1$, 因此

$$P(z)=z^n+o(z^n), \text{当} |z|\to\infty \text{时}. \tag{33}$$

我们记得 (见例 15) 在映射 $z\mapsto z^n$ 之下, 中心在点 O 半径为 r 的圆周变成中心在点 O 半径为 r^n 的圆周, 那么, 对于足够大的 r 值, 根据 (33), 圆周 $|z|=r$ 的像将以甚小之相对误差重合于 w 平面上的圆周 $|w|=r^n$(图 44). 在任何情况下, 像都是围住点 $w=0$ 的一条曲线, 这一点是重要的.

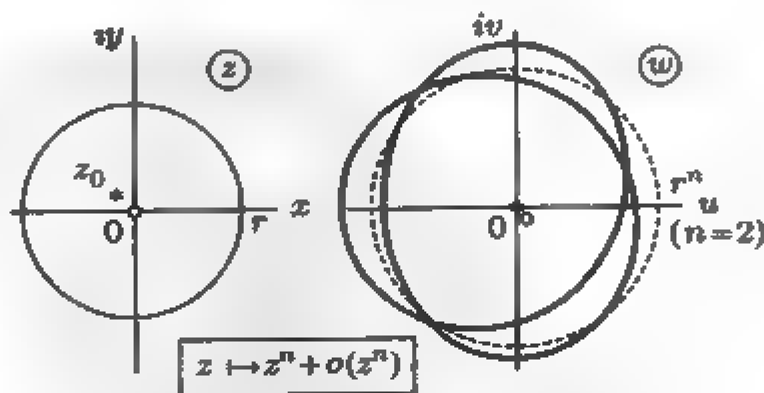


图 44

如果把圆 $|z|\leqslant r$ 看作是一个张在圆周 $|z|=r$ 上的薄膜, 那么, 在由多项式 $w=P(z)$ 实现的连续映射下, 这个薄膜变成张在圆周的像上面的薄膜. 但由于这个

像围住了点 $w=0$, 所以这个薄膜必有某点与 $w=0$ 重合, 这就意味着, 在圆 $|z|<r$ 内存在一点 z_0, 它在映射 $w=P(z)$ 下正好变成 $w=0$, 即 $P(z_0)=0$.

这个完全直观的思想导致一系列非常重要非常有用的拓扑概念 (关于点的路径指标, 映射的级), 借助这些概念就可以做出完全的证明, 其正确性尽可能地不仅是对于多项式才成立. 不过很遗憾, 这些考察会使我们偏离我们当前所研究的基本课题; 因此我们将基于已经足够熟悉的思路做出另一个证明.

定理 2 每个次数 $n \geqslant 1$ 的复系数多项式

$$P(z)=c_0+\cdots+c_nz^n$$

在$\mathbb{C}$中都有根.

◀不损伤定理结论的一般性, 显然可以认为 $c_n=1$. 设 $\mu=\inf\limits_{z\in\mathbb{C}}|P(z)|$. 由于

$$P(z)=z^n\left[1+\frac{c_{n-1}}{z}+\cdots+\frac{c_0}{z^n}\right],$$

所以

$$|P(z)|\geqslant|z|^n\left[1-\left(\frac{|c_{n-1}|}{|z|}+\cdots+\frac{|c_0|}{|z|^n}\right)\right].$$

于是显然, 对于足够大的值 R, 当 $|z|>R$ 时, $|P(z)|>\max\{1,2\mu\}$. 因此, 使

$$0<|P(z_k)|-\mu<\frac{1}{k}$$

的一列点 $z_2,z_3,\cdots,z_k,\cdots$ 落在圆 $|z|\leqslant R$ 中.

我们来检验, 在 $\mathbb{C}$ 中 (甚至在圆 $|z|\leqslant R$ 中) 有使 $|P(z_0)|=\mu$ 的点 z_0. 为此指出, 若 $z_k=x_k+iy_k$, 则

$$\max\{|x_k|,|y_k|\}\leqslant|z_k|\leqslant R,$$

从而实数列 $\{x_k\}$、$\{y_k\}$ 都是有界的. 先从 $\{x_k\}$ 中抽出收敛的子列 $\{x_{k_l}\}$, 再从 $\{y_{k_l}\}$ 中抽出收敛的子列 $\{y_{k_{l_m}}\}$, 就得到 $\{z_k\}$ 的子列

$$z_{k_{l_m}}=x_{k_{l_m}}+iy_{k_{l_m}},$$

它有极限

$$\lim_{m\to\infty}z_{k_{l_m}}=\lim_{m\to\infty}x_{k_{l_m}}+i\lim_{m\to\infty}y_{k_{l_m}}=x_0+iy_0=z_0,$$

且由于当 $m\to\infty$ 时 $|z_{k_{l_m}}|\to|z_0|$, 所以 $|z_0|\leqslant R$. 为避免写法累赘, 我们不转到子列, 而认为数列 $\{z_k\}$ 本身即已收敛. 从 $P(z)$ 在 $z_0\in\mathbb{C}$ 的连续性推出

$$\lim_{k\to\infty}P(z_k)=P(z_0).$$

但这样一来[①], $|P(z_0)| = \lim_{k\to\infty} |P(z_k)| = \mu$.

现在假定 $\mu > 0$, 我们来引出矛盾. 若 $P(z_0) \neq 0$, 则考察多项式 $Q(z) = \frac{P(z+z_0)}{P(z_0)}$ 按其构成, $Q(0) = 1$, 且 $|Q(z)| = \frac{|P(z+z_0)|}{|P(z_0)|} \geqslant 1$.

由于 $Q(0) = 1$, 多项式 $Q(z)$ 有如下形式

$$Q(z) = 1 + q_k z^k + q_{k+1} z^{k+1} + \cdots + q_n z^n,$$

其中 $|q_k| \neq 0$ 且 $1 \leqslant k \leqslant n$. 若 $q_k = \rho e^{i\psi}$, 则对于 $\varphi = \frac{\pi - \psi}{k}$ 有

$$q_k \cdot (e^{i\varphi})^k = \rho e^{i\psi} \cdot e^{i(\pi-\psi)} = \rho e^{i\pi} = -\rho = -|q_k|.$$

于是当 $z = re^{i\varphi}$ 且 r 充分接近于零时, 得到

$$\begin{aligned} |Q(re^{i\varphi})| &\leqslant |1 + q_k z^k| + (|q_{k+1} z^{k+1}| + \cdots + |q_n z^n|) \\ &= |1 - r^k |q_k|| + r^{k+1}(|q_{k+1}| + \cdots + |q_n| r^{n-k-1}) \\ &= 1 - r^k(|q_k| - r|q_{k+1}| - \cdots - r^{n-k}|q_n|) < 1. \end{aligned}$$

但是当 $z \in \mathbb{C}$ 时 $|Q(z)| \geqslant 1$. 所得的矛盾表明 $P(z_0) = 0$. ►

注 1 关于任意复系数代数方程在 $\mathbb{C}$ 中都可解的定理 (传统上常称之为代数基本定理) 的第一个证明是由高斯给出的. 他一向注意把所谓 “虚数” 引向完全现实的生活, 为它找到了各式各样的深刻的应用.

注 2 我们知道, 实系数多项式 $P(z) = a_0 + \cdots + a_n z^n$ 并不总有实根. 但与任意的复系数多项式相比它具有这样的特性, 若 $P(z_0) = 0$, 则亦有 $P(\bar{z}_0) = 0$. 实际上, 从共轭数的定义及复数的加法法则推出 $\overline{(z_1 + z_2)} = \bar{z}_1 + \bar{z}_2$. 从复数的三角式的写法及复数的乘法法则见到

$$\begin{aligned} \overline{(z_1 \cdot z_2)} &= \overline{(r_1 e^{i\varphi_1} \cdot r_2 e^{i\varphi_2})} = \overline{r_1 r_2 e^{i(\varphi_1+\varphi_2)}} = r_1 r_2 e^{-i(\varphi_1+\varphi_2)} \\ &= r_1 e^{-i\varphi_1} \cdot r_2 e^{-i\varphi_2} = \bar{z}_1 \cdot \bar{z}_2. \end{aligned}$$

因此

$$\overline{P(z_0)} = \overline{a_0 + \cdots + a_n z_0^n} = \bar{a}_0 + \cdots + \bar{a}_n \bar{z}_0^n = a_0 + \cdots + a_n \bar{z}_0^n = P(\bar{z}_0),$$

从而若 $P(z_0) = 0$, 则 $\overline{P(z_0)} = P(\bar{z}_0) = 0$.

---

①注意, 我们一方面证明了, 从任何依模有界的复数列中都可抽出收敛的子列, 另一方面也展示了一个可以用来证明关于闭区间上连续函数最小值定理的方法, 在现在的场合, 这是在圆 $|z| \leqslant R$ 中进行的.

推论 1 任何一个次数 $n \geqslant 1$ 的复系数多项式 $P(z) = c_0 + \cdots + c_n z^n$ 都可以表示成

$$P(z) = c_n(z - z_1) \cdots (z - z_n), \tag{34}$$

其中 z_1、$\cdots$、$z_n \in \mathbb{C}$(z_1、$\cdots$、z_n 不必互不相同), 并且若不计较因子的次序, 则表达式 (34) 是唯一的.

◀ 根据多项式 $P(z)$ 除以次数 $m \leqslant n$ 的多项式 $Q(z)$ 的除法法则 ("欧几里得辗转相除法"), 求得 $P(z) = q(z)Q(z) + r(z)$, 其中 $q(z)$ 和 $r(z)$ 是多项式, 且 $r(z)$ 的次数小于 $Q(z)$ 的次数, 即小于 m. 于是, 若 $m = 1$, 则 $r(z) = r$ 是常数.

设 z_1 是多项式 $P(z)$ 的根. 那么 $P(z) = (z - z_1)q(z) + r$, 且由于 $P(z_1) = r$, 所以 $r = 0$. 就是说, 若 z_1 是多项式 $P(z)$ 的根, 则成立着表达式 $P(z) = (z - z_1)q(z)$. 多项式 $q(z)$ 的次数等于 $n - 1$, 并且对它可以重复同样的讨论. 根据归纳法我们得到 $P(z) = c(z - z_1) \cdots (z - z_n)$. 由于应有 $cz^n = c_n z^n$, 所以 $c = c_n$. ▶

推论 2 任何实系数多项式

$$P(z) = a_0 + \cdots + a_n z^n$$

都可以展开成一次和二次实系数多项式的乘积.

◀ 此命题从上述推论 1 和注 2 推出. 根据注 2, 与 $P(z)$ 的根 z_k 一起, $\bar{z}_k$ 也是 $P(z)$ 的根[①]. 于是在展开式 (34) 中把因子 $(z - z_k)$ 和 $(z - \bar{z}_k)$ 乘起来, 得到实系数二次多项式

$$z^2 + (-z_k - \bar{z}_k)z + |z_k|^2.$$

数 c_n 此时等于 a_n, 是实数, 因此可以把它乘到展开式的一个括号里面去而不改变此因子的次数. ▶

把展开式 (34) 中相同的因子乘在一起, 可以把它改写成

$$P(z) = c(z - z_1)^{k_1} \cdots (z - z_p)^{k_p}. \tag{35}$$

数 k_j 叫做根 z_j 的重数.

由于 $P(z) = (z - z_j)^{k_j} Q(z)$, 其中 $Q(z_j) \neq 0$, 所以

$$P'(z) = k_j(z - z_j)^{k_j - 1} Q(z) + (z - z_j)^{k_j} Q'(z) = (z - z_j)^{k_j - 1} R(z),$$

其中 $R(z_j) - k_j Q(z_j) \neq 0$. 这样我们就得到下面的结论.

推论 3 多项式 $P(z)$ 的每个重数为 $k_j > 1$ 的根 z_j 都是 $P(z)$ 的导数 — 多项式 $P'(z)$ 的 $k_j - 1$ 重根.

①需要指出, 对于 $z_k \notin \mathbb{R}$, 作为根, z_k 与 $\bar{z}_k$ 有相同的重数. —— 译者注

根据这个推论和展开式 (35), 我们不用求多项式 $P(z)$ 的根就可以求出形如 $(z-z_1)\cdots(z-z_p)$ 的多项式 $p(z)$, 它的根与 $P(z)$ 的根一样, 但都是一重的.

实际上, 按照欧几里得算法, 先求出 $P(z)$ 和 $P'(z)$ 的最大公因子 —— 多项式 $q(z)$. 根据推论 3、展开式 (35) 以及定理 2, 多项式 $q(z)$ 与乘积 $(z-z_1)^{k_1-1}\cdots(z-z_p)^{k_p-1}$ 顶多差一个常数因子, 因此, 用 $q(z)$ 去除 $P(z)$, 再除以 z^p 项的系数就得到多项式 $p(z)=(z-z_1)\cdots(z-z_p)$.

现在考察两个多项式的比 $R(x)=\dfrac{P(x)}{Q(x)}$, 其中 $Q(x)\neq\text{const}$. 如果 $P(x)$ 的次数大于 $Q(x)$ 的次数, 那么, 用多项式的除法律, 把 $P(x)$ 表成

$$P(x)=p(x)\cdot Q(x)+r(x),$$

其中 $p(x)$ 和 $r(x)$ 是多项式且 $r(x)$ 的次数小于 $Q(x)$ 的次数. 这样一来, 得到了 $R(x)$ 的表达式

$$R(x)=p(x)+\frac{r(x)}{Q(x)},$$

其中分式 $\dfrac{r(x)}{Q(x)}$ 依分子的次数小于分母的次数这个意义来说已是真分式了.

现在我们要叙述这样一个推论, 它是关于把真分式表示成所谓最简分式的和的.

推论 4 a) *若* $Q(z)=(z-z_1)^{k_1}\cdots(z-z_p)^{k_p}$ *且* $\dfrac{P(z)}{Q(z)}$ *是真分式, 则对于* $\dfrac{P(z)}{Q(z)}$ *存在唯一的下述形式的表示式*

$$\frac{P(z)}{Q(z)}=\sum_{j=1}^{p}\left(\sum_{k=1}^{k_j}\frac{a_{jk}}{(z-z_j)^k}\right). \tag{36}$$

b) *若* $P(x)$ *和* $Q(x)$ *是实系数多项式且*

$$Q(x)=(x-x_1)^{k_1}\cdots(x-x_l)^{k_l}(x^2+p_1x+q_1)^{m_1}\cdots(x^2+p_nx+q_n)^{m_n},$$

则对于真分式 $\dfrac{P(x)}{Q(x)}$ *存在唯一的下述形式的表示式*

$$\frac{P(x)}{Q(x)}=\sum_{j=1}^{l}\left(\sum_{k=1}^{k_j}\frac{a_{jk}}{(x-x_j)^k}\right)+\sum_{j=1}^{n}\left(\sum_{k=1}^{m_j}\frac{b_{jk}x+c_{jk}}{(x^2+p_jx+q_j)^k}\right), \tag{37}$$

其中 a_{jk}、b_{jk}、c_{jk} *都是实数.*

我们指出, 实际寻求展开式 (36) 或 (37) 的万能的虽然并不总是最简单的方法是待定系数法. 这个方法是这样的, 把 (36) 或 (37) 右端的和用公分母 $Q(x)$ 来通分, 然后将所得的分子的系数与多项式 $P(x)$ 的相应的系数来对比. 此时我们得到一个线性方程组, 根据推论 4, 这个方程组永远可以单值解出.

因为, 通常我们只对于具体的分式的展开式感兴趣, 我们用待定系数法来求出这个展开式, 所以除了根据推论 4 确信此事总可做到之外, 我们暂且并不需要别的什么. 由于这个缘故, 我们不打算对推论 4 进行证明. 它的证明通常以代数的语言载于高等代数教科书中或以分析的语言载于复变函数论的教科书中.

我们考察一个专门挑选的例子, 通过它对上面的叙述作出说明.

例 17 设

$$P(x) = 2x^6 + 3x^5 + 6x^4 + 6x^3 + 10x^2 + 3x + 2,$$
$$Q(x) = x^7 + 3x^6 + 5x^5 + 7x^4 + 7x^3 + 5x^2 + 3x + 1;$$

要求算出分式 $\dfrac{P(x)}{Q(x)}$ 的展开式 (37).

首先, 麻烦的是我们不知道多项式 $Q(x)$ 的展开式. 我们试着把情况简化, 避开 $Q(x)$ 可能会有的重根. 求出

$$Q'(x) = 7x^6 + 18x^5 + 25x^4 + 28x^3 + 21x^2 + 10x + 3.$$

按欧几里得算法进行相当冗繁但可完成的计算, 求出多项式 $Q(x)$ 和 $Q'(x)$ 的最大公因子

$$d(x) = x^4 + 2x^3 + 2x^2 + 2x + 1.$$

我们写出的是 x 的最高次幂的系数为 1 的最大公因子.

用 $d(x)$ 除 $Q(x)$ 得多项式

$$q(x) = x^3 + x^2 + x + 1,$$

它与多项式 $Q(x)$ 有同样的根, 不过都是一重的. 容易猜出 -1 是多项式 $q(x)$ 的根. 用 $x+1$ 除 $q(x)$ 后得到 x^2+1. 因此

$$q(x) = (x+1)(x^2+1).$$

逐次用 x^2+1 和 $x+1$ 去除 $d(x)$, 而后求得 $d(x)$ 的展开式

$$d(x) = (x+1)^2(x^2+1),$$

随后就得到展开式

$$Q(x) = (x+1)^3(x^2+1)^2.$$

于是, 根据推论 4b) 求得分式 $\dfrac{P(x)}{Q(x)}$ 的如下形式的展开式:

$$\frac{P(x)}{Q(x)} = \frac{a_{11}}{x+1} + \frac{a_{12}}{(x+1)^2} + \frac{a_{13}}{(x+1)^3} + \frac{b_{11}x + c_{11}}{x^2+1} + \frac{b_{12}x + c_{12}}{(x^2+1)^2}.$$

将右端通分并比较分子中所得多项式的系数和多项式 $P(x)$ 的相应的系数, 得到七个七元方程, 解此方程组, 最后得到

$$\frac{P(x)}{Q(x)}=\frac{1}{x+1}-\frac{2}{(x+1)^2}+\frac{1}{(x+1)^3}+\frac{x-1}{x^2+1}+\frac{x+1}{(x^2+1)^2}.$$

练 习

1. 使用复数的几何解释:

a) 解释不等式 $|z_1+z_2|\leqslant|z_1|+|z_2|$ 和 $|z_1+\cdots+z_n|\leqslant|z_1|+\cdots+|z_n|$;

b) 指出平面 $\mathbb{C}$ 上满足关系式 $|z-1|+|z+1|\leqslant 3$ 的点的几何位置;

c) 表出数 1 的所有的 n 次根并求它们的和;

d) 解释由公式 $z\mapsto\bar{z}$ 给出的平面 $\mathbb{C}$ 的变换的效用.

2. 求和

a) $1+q+\cdots+q^n$;

b) $1+q+\cdots+q^n+\cdots$, 当 $|q|<1$ 时;

c) $1+e^{i\varphi}+\cdots+e^{in\varphi}$;

d) $1+re^{i\varphi}+\cdots+r^ne^{in\varphi}$;

e) $1+re^{i\varphi}+\cdots+r^ne^{in\varphi}+\cdots$, 当 $|r|<1$ 时;

f) $1+r\cos\varphi+\cdots+r^n\cos n\varphi$;

g) $1+r\cos\varphi+\cdots+r^n\cos n\varphi+\cdots$, 当 $|r|<1$ 时;

h) $1+r\sin\varphi+\cdots+r^n\sin n\varphi$;

i) $1+r\sin\varphi+\cdots+r^n\sin n\varphi+\cdots$, 当 $|r|<1$ 时.

3. 求出复数 $\lim\limits_{n\to\infty}\left(1+\frac{z}{n}\right)^n$ 的模和辐角并证实这个数是 e^z.

4. a) 证明关于 w 的方程 $e^w=z$ 有解 $w=\ln|z|+i\mathrm{Arg}\,z$. 自然认为 w 是数 z 的自然对数. 因此 $w=\ln z$ 不是函数关系式, 因为 $\mathrm{Arg}\,z$ 是多值的.

b) 求出 $\ln 1$ 和 $\ln i$.

c) 置 $z^\alpha=e^{\alpha\ln z}$. 求 1^π 和 i^i.

d) 使用表达式 $w=\sin z=\frac{1}{2i}(e^{iz}-e^{-iz})$, 求出对于 $z=\arcsin w$ 的表达式.

e) 在 $\mathbb{C}$ 中有没有点能使 $|\sin z|=2$?

5. a) 研究一下是不是在平面 $\mathbb{C}$ 的所有的点处函数 $f(z)=\frac{1}{1+z^2}$ 都连续.

b) 在 $z_0=0$ 处把函数 $\frac{1}{1+z^2}$ 展开成幂级数并求它的收敛半径.

c) 对函数 $\frac{1}{1+\lambda^2z^2}$ 求解问题 a) 和 b) 其中 $\lambda\in\mathbb{R}$ 是参数. 关于平面 $\mathbb{C}$ 上什么样的点的相互位置决定收敛半径, 你有没有什么想法? 当限制在实轴上关于 $x\in\mathbb{R}$ 来展开函数 $f(x)=\frac{1}{1+\lambda^2x^2}(\lambda\in\mathbb{R})$ 时, 此事能明白否?

6. a) 研究一下柯西函数

$$f(z)=\begin{cases}e^{-\frac{1}{z^2}}, & z\neq 0,\\ 0, & z=0\end{cases}$$

在点 $z=0$ 处是否连续.

b) 问题 a) 中的函数 f 在实轴上的限制 $f|_{\mathbb{R}}$ 是否连续?

c) a) 中的函数 f 在点 $z_0=0$ 处有泰勒级数吗?

d) 有没有在 $z_0 \in \mathbb{C}$ 解析的函数, 它的泰勒级数只在点 z_0 处收敛?

e) 想出一个只在点 z_0 处收敛的幂级数 $\sum\limits_{n=0}^{n} c_n(z-z_0)^n$.

7. a) 在幂级数 $\sum\limits_{n=0}^{\infty} A_n(z-a)^n$ 中作形式的代换 $z-a=(z-z_0)+(z_0-a)$ 并合并同类项, 将得到一个级数 $\sum\limits_{n=0}^{\infty} c_n(z-z_0)^n$, 试用量 $A_k, (z_0-a)^k, k=0,1,\cdots$ 表示这个级数的系数.

b) 试验证, 如果原级数在圆 $|z-a|<k$ 中收敛, 而 $|z_0-a|=r<\mathbb{R}$, 则定义 $c_n(n=0,1,\cdots)$ 的那些级数绝对收敛, 而且, 级数 $\sum\limits_{n=0}^{\infty} c_n(z-z_0)^n$ 当 $|z-z_0|<\mathbb{R}-r$ 时收敛.

c) 试证, 如果在圆 $|z-a|<\mathbb{R}$ 中有 $f(z)=\sum\limits_{n=0}^{\infty} A_n(z-a)^n$, 而 $|z_0-a|<\mathbb{R}$, 那么, 在圆 $|z-z_0|<\mathbb{R}-|z_0-a|$ 中函数 f 有幂级数表示 $f(z)=\sum\limits_{n=0}^{\infty} c_n(z-z_0)^n$.

8. 验证:

a) 当点 $z \in \mathbb{C}$ 跑遍圆周 $|z|=r>1$, 点 $w=z+z^{-1}$ 跑遍以 0 为中心、点 ± 2 为焦点的椭圆周;

b) 当作复数平方时, 更准确地说, 在映射 $w \mapsto w^2$ 下这样的椭圆周将变成以 0 为焦点的跑遍两次的椭圆周

c) 在映射 $w \mapsto w^2$ 下, 任何以零为中心的椭圆周将变成以 0 为焦点的椭圆周.

§6. 自然科学中应用微分学的一些例子

这节我们将分析自然科学中的几个问题. 这些问题的出处, 彼此相差甚远, 但我们会看到它们却有相当接近的数学模型. 这个模型不是别的, 乃是关于我们在问题中所感兴趣的函数的最简单的微分方程. 顺便提及, 我们一般正是从研究一个这样的例子 —— 二体问题而开始去建立微分学的. 这个问题归结为一个方程组, 目前我们还不能去研究它. 这里我们只考察那些按我们现有的水平能够彻底解决的问题. 从这节的一系列例子当中, 我们除了见到数学工具的具体运用而满足之外, 特别地, 会更加相信, 不但指数函数 $\exp x$ 的产生是十分自然的, 而且它向复数域的推广也是非常有效的.

1. 齐奥尔柯夫斯基公式 考察远离有引力的物体, 在开阔宇宙中直线运动的火箭 (图 45).

设 $M(t)$ 是火箭 (连同燃料) 在时刻 t 的质量; $V(t)$ 是它在时刻 t 的速度; ω 是燃料爆燃由火箭的管咀喷出的 (相对于火箭的) 速度.

我们来确定这些量之间的关系.

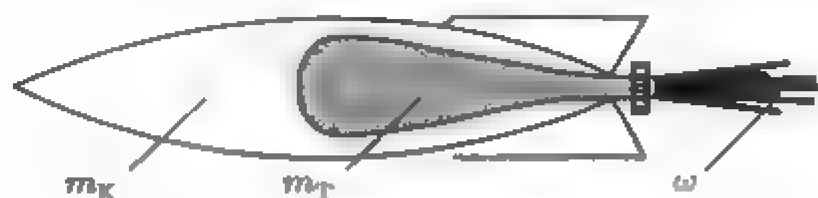


图 45

根据所做的假定, 火箭连同燃料可以看作封闭系统, 因此它的冲量 (或动量) 不随时间变化.

在时刻 t 系统的冲量等于 $M(t)\cdot V(t)$.

在时刻 $t+h$ 火箭连同它所剩下的燃料的冲量等于 $M(t+h)\cdot V(t+h)$, 而在这段时间内喷出的燃料质量是

$$|\Delta M| = |M(t+h) - M(t)| = -(M(t+h) - M(t)),$$

喷出的这些燃料的冲量 ΔI 由下式界定

$$(V(t)-\omega)|\Delta M| < \Delta I < (V(t+h)-\omega)|\Delta M|,$$

即

$$\Delta I = (V(t)-\omega))|\Delta M| + \alpha(h)|\Delta M|,$$

且由 $V(t)$ 的连续性推出, 当 $h\to 0$ 时, $\alpha(h)\to 0$.

对比系统在时刻 t 和 $t+h$ 的冲量, 有

$$M(t)V(t) = M(t+h)V(t+h) + (V(t)-\omega)|\Delta M| + \alpha(h)|\Delta M|.$$

代入

$$|\Delta M| = -(M(t+h) - M(t))$$

并化简, 得

$$M(t+h)(V(t+h)-V(t)) = -\omega(M(t+h)-M(t)) + \alpha(h)(M(t+h)-M(t)).$$

以 h 除此式并令 $h\to 0$ 取极限, 得到

$$M(t)V'(t) = -\omega M'(t). \tag{1}$$

这就是要找的我们所感兴趣的函数 $V(t)$、$M(t)$ 和它们的导数之间的关系式.

现在必须从函数的导数之间的关系式出发来求出函数 $V(t)$、$M(t)$ 本身之间的关系. 一般说来, 这类问题比从函数之间的已知关系求导数之间的关系要困难得多. 然而在我们的情况下问题倒是可以完全用初等办法来解决.

实际上, 等式 (1) 除以 $M(t)$ 之后可改写成

$$V'(t) = (-\omega \ln M)'(t). \tag{2}$$

但若两个函数的导数在某个区间上相同, 则在此区间上函数本身只不过相差某个常数.

于是从 (2) 推出

$$V(t) = -\omega \ln M(t) + c. \tag{3}$$

如果知道, 例如说, $V(0) = V_0$, 那么, 这个条件就完全确定了常数 c. 实际上, 从 (3) 求出

$$c = V_0 + \omega \ln M(0),$$

从而求出要找的公式①

$$V(t) = V_0 + \omega \ln \frac{M(0)}{M(t)} \tag{4}$$

指出下述关系是有用的. 若 m_k 是火箭本身的质量, m_T 是燃料的质量, 而 V 是燃料完全烧尽时火箭达到的终极速度, 那么, 把 $M(0) = m_k + m_T$ 和 $M(t) = m_k$ 代入 (4) 就求得

$$V = V_0 + \omega \ln \left(1 + \frac{m_T}{m_k}\right).$$

这个公式特别清楚地表明, 终极速度与其说决定于对数符号下的比值 $\frac{m_T}{m_k}$, 不如说决定于由燃料类型确定的喷出速度 ω. 特别地, 由这个公式推出, 若 $V_0 = 0$, 则要使本身质量为 m_k 的火箭得到速度 V, 必须开始时带够如下数量的燃料:

$$m_T = m_k(e^{\frac{V}{\omega}} - 1).$$

2. 气压公式　这是一个表示大气压强与海拔高度关系的公式.

设 $p(h)$ 是高 h 处的压强. 因为 $p(h)$ 是在高 h 处一平方厘米面积上面的空气柱的重量, 所以 $p(h+\Delta)$ 与 $p(h)$ 相差的是高 h 处一平方厘米大小的平面和它上方高 $h+\Delta$ 处的同样的平面之间空气柱的重量. 设 $\rho(h)$ 是空气在高度 h 处的密度. 因为 $\rho(h)$ 连续地依赖于 h, 所以, 可以认为所说的空气柱的质量可由下式计算:

$$\rho(\xi) \cdot \mathrm{g/cm^3} \cdot 1\mathrm{cm^2} \cdot \Delta\mathrm{cm} = \rho(\xi) \cdot \Delta\ \mathrm{g}$$

其中 ξ 是由 h 到 $h+\Delta$ 之间的某个高度水平的数目. 就是说, 这个空气柱的重量是 $g \cdot \rho(\xi)\Delta$②.

①这个公式有时冠以齐奥尔柯夫斯基 (К. Э. Циолковский) (1857—1935) 的名字, 他是俄罗斯学者, 宇宙飞行理论的奠基人. 但是, 第一个得到这个公式的人大概是俄罗斯力学家米歇尔斯基 (И. В. Мещерский) (1859—1935), 发表在 1897 年他的关于变质量质点动力学的文章中.

②在大气明显存在的范围内, 可认为重力加速度 g 是常数.

于是

$$p(h+\Delta)-p(h)=-g\rho(\xi)\Delta.$$

以 Δ 除此式两端且令 $\Delta\to 0$ 取极限, 注意此时也有 $\xi\to h$, 得到

$$p'(h)=-g\cdot\rho(h). \tag{5}$$

这样一来, 压强变化的速度显然是与相应高度处的空气密度成比例的.

为了得到关于函数 $p(h)$ 的方程, 我们从 (5) 中消去函数 $\rho(h)$. 根据克拉贝龙[1]定律, 气体的压强 p、克分子体积 V 和开氏[2] 温度 T 由下式相联系

$$\frac{pV}{T}=R, \tag{6}$$

其中 R 是所谓万用气体常数. 若 M 是一克分子空气的质量, 而 V 是它的体积, 则 $\rho=\frac{M}{V}$. 因此, 从 (6) 求出

$$p=\frac{1}{V}RT=\frac{M}{V}\frac{R}{M}\cdot T=\rho\frac{R}{M}T.$$

令 $\lambda=\frac{R}{M}T$, 则有

$$p=\lambda(T)\rho. \tag{7}$$

现在如果认为我们所观察的空气层的温度是常数, 那么从 (5) 和 (7) 最后求得

$$p'(h)=-\frac{g}{\lambda}p(h). \tag{8}$$

这个微分方程可以改写成

$$\frac{p'(h)}{p(h)}=-\frac{g}{\lambda},$$

即

$$(\ln p)'(h)=\left(-\frac{g}{\lambda}h\right)'.$$

由此得

$$\ln p(h)=-\frac{g}{\lambda}h+c,$$

即

$$p(h)=e^{c}\cdot e^{-\frac{g}{\lambda}h}.$$

因子 e^{c} 可由已知的初始条件 $p(0)=p_0$ 确定, 据此 $e^{c}=p_0$.

这样, 我们就求出了如下的压强对高度的依赖关系:

$$p=p_0e^{-\frac{g}{\lambda}h}. \tag{9}$$

---

[1] 克拉贝龙 (Clapeyron)(1799—1864) —— 法国物理学家, 从事热力学的研究.

[2] 汤姆森 (W. Thomson)(1824—1907) —— 著名英国物理学家, 后称 Lord Kelvin(开尔文勋爵).

对于室温 (300° K – 27 °C 左右) 的空气, 知道值 $\lambda \approx 7.7 \times 10^8 (\mathrm{cm/s})^2$. 还知道 $g \approx 10^3 \mathrm{cm/s}^2$ 代入 g 和 λ 的这些数值之后, 公式 (9) 就得到了完全终极的形式. 特别地, 从 (9) 见到, 在高度 $h = \dfrac{\lambda}{g} = 7.7 \times 10^5$ cm $= 7.7$ km 处, 压强将下降到初始压强的 $e^{-1}\left(\approx \dfrac{1}{3}\right)$. 若落到深 7.7 km 左右的竖井底处, 则气压要增大到 $e(\approx 3)$ 倍.

3. 放射衰变、连锁反应及原子反应堆 众所周知, 重元素 (主要是超铀元素) 的原子核时时进行着自发的衰变. 这就是所谓的*自然辐射*.

辐射的基本统计规律 (因是统计规律, 所以对量不是太少的高浓度物质是真确的) 告诉我们, 物质由时刻 t 经过一小段时间 h 所衰减的量与 h 以及直到时刻 t 尚未衰变的原子的总量 $N(t)$ 成正比, 即

$$N(t+h) - N(t) \approx -\lambda N(t) \cdot h,$$

其中 $\lambda > 0$ 是一个表征所说的化学元素的数值系数.

于是, 函数 $N(t)$ 满足已经熟知的微分方程

$$N'(t) = -\lambda N(t), \tag{10}$$

由此推出

$$N(t) = N_0 e^{-\lambda t},$$

其中 $N_0 = N(0)$ 是物质原子的初始总量.

从最初的原子数量衰减一半所经历的时间 T 叫做*半衰期*. 因此可从方程 $e^{-\lambda T} = \dfrac{1}{2}$ 求得量 T, 即 $T = \dfrac{\ln 2}{\lambda} \approx \dfrac{0.69}{\lambda}$.

例如, 对于钋 $\mathrm{Po}^{210}, T \approx 138$ 天; 对于镭 $\mathrm{Ra}^{226}, T \approx 1600$ 年; 对于铀 $\mathrm{U}^{235}, T \approx 7.1 \times 10^8$ 年, 而对于铀的同位素 $\mathrm{U}^{238}, T \approx 4.5 \times 10^9$ 年.

*核反应*指的是核之间的相互作用或者核与基本粒子的相互作用, 其结果出现新类型的核. 这可以是核的聚变, 即一些较轻的元素的核合成较重的元素的核 (例如两个重氢的核伴随损失质量和释放能量合成氦的核); 也可以是一个核衰变而形成一个或多个较轻的元素的核. 尤其是当中子与铀 U^{235} 的核相撞的时候, 大约在一半的情况下要发生这样的衰变. 当铀核分裂时, 形成二三个新的中子, 它们能继续参与与核的相互作用, 引起核的分裂, 因而使中子增殖. 这种类型的核反应叫做*连锁反应*.

我们来描述放射性物质连锁反应的基本数学模型, 求出中子数目 $N(t)$ 随时间变化的规律.

取半径为 r 的球形物体. 如果 r 不太小的话, 则在由时刻 t 算起的时间间隔 h 中, 一方面要产生数量正比于 h 和 $N(t)$ 的新中子, 另一方面要失去一部分跑到球外的中子.

如果 v 是中子的速度, 那么, 能在时间 h 后离开球体的中子只能是那些处在离球面距离不超过 vh 的区域当中且速度方向基本上是径向的那些中子. 如果认为这个区域中经过时间 h 后跑到球外的中子数与此区域中的中子总数的比不随时刻 t 的变化而变化, 并且认为球中的中子是均匀分布的, 那么就可以说, 经过时间 h 后失去的中子数正比于 $N(t)$, 也正比于上述靠近边界的区域的体积与球的体积的比.

由上所述得到等式

$$N(t+h)-N(t)\approx \alpha N(t)h-\frac{\beta}{r}N(t)h, \tag{11}$$

$\left(\text{因为所考虑的区域的体积约等于}4\pi r^2vh,\text{而球的体积是}\frac{4}{3}\pi r^3\right)$. 系数 α 和 β 仅与所考虑的放射性物质有关.

将关系式 (11) 除以 h 再令 $h\to 0$ 取极限, 得到

$$N'(t)=\left(\alpha-\frac{\beta}{r}\right)N(t). \tag{12}$$

由此

$$N(t)=N_0e^{\left(\alpha-\frac{\beta}{r}\right)t}.$$

从所得的公式见到, 当 $\left(\alpha-\frac{\beta}{r}\right)>0$ 时, 中子数目依指数律随着时间增长. 这种增长的特点是这样的, 不管初始条件 N_0 如何, 在很短的时间内就会使物体实际上完全衰变, 并伴随着释放出极大的能量 —— 这就是爆炸.

如果 $\left(\alpha-\frac{\beta}{r}\right)<0$, 那么反应很快就要终止, 因为失掉的中子比产生的多.

如果成立界于上边考虑的两种情形之间的条件 $\alpha-\frac{\beta}{r}=0$, 则反应中新生的中子和退出反应的中子达到平衡, 结果中子数目大致不变.

使 $\alpha-\frac{\beta}{r}=0$ 成立的量 r 叫做临界半径, 而具有临界半径的球状物体的质量叫做所说物质的临界质量.

对于铀 U^{235}, 临界半径约等于 8.5 cm, 而临界质量接近 50 kg.

在依靠放射性物质的连锁反应来预热蒸汽的原子反应堆中, 设置有人工中子源, 它在单位时间内把确定的 n 个中子送进裂变的物质. 因此, 对于原子反应堆, 方程式 (12) 要稍加改变:

$$N'(t)=\left(\alpha-\frac{\beta}{r}\right)N(t)+n. \tag{13}$$

可以用解方程 (12) 的同样办法来解这个方程, 因为当 $\alpha-\frac{\beta}{r}\neq 0$ 时,

$$\frac{N'(t)}{\left(\alpha-\frac{\beta}{r}\right)N(t)+n}$$

是函数 $\frac{1}{\alpha-\frac{\beta}{r}}\times\ln\left[\left(\alpha-\frac{\beta}{r}\right)N(t)+n\right]$ 的导数. 方程 (13) 的解是

$$N(t)=\begin{cases}N_0e^{\left(\alpha-\frac{\beta}{r}\right)t}-\frac{n}{\alpha-\frac{\beta}{r}}\{1-e^{\left(\alpha-\frac{\beta}{r}\right)t}\}, & \text{当}\alpha-\frac{\beta}{r}\neq 0,\\ N_0+nt, & \text{当}\alpha-\frac{\beta}{r}=0.\end{cases}$$

由这个解看出, 若 $\alpha-\frac{\beta}{r}>0$(超过临界质量), 则发生爆炸. 若不足临界质量, 即 $\alpha-\frac{\beta}{r}<0$, 则很快就有

$$N(t)\approx\frac{n}{\frac{\beta}{r}-\alpha}.$$

因此, 如果把放射性物质的质量维持在不足而又接近临界的状态, 那么不管补充中子源的强弱如何, 即不管 n 的值多大, 都可以得到很大的值 $N(t)$, 这就意味着得到很大的反应强度. 把过程保持在接近临界状态是一件奥妙且要由相当复杂的自动控制系统来实现的事情.

4. 空气中的落体 现在我们来研究物体在重力作用下落向地球的速度.

如果不存在空气阻力的话, 那么, 从相对不大的高度落下时, 我们会有关系式

$$\dot{v}(t)=g. \tag{14}$$

这是从牛顿第二定律$ma=F$ 和万有引力定律推出的. 根据万有引力定律, 当 $h\ll R$ 时 (R 是地球的半径)

$$F(t)=G\frac{M\cdot m}{(R+h(t))^2}\approx G\frac{M\cdot m}{R^2}=g\cdot m.$$

在空气中运动的物体受到与运动速度有关的阻力, 其结果使得重物在空气中自由下落的速度不是无限地增加而是稳定在某个水平上. 例如, 做延迟跳伞时, 跳伞者的速度在最下层空气中稳定在 50m 每秒至 60m 每秒的范围内.

在速度由 0 到 80m 每秒的范围内, 我们认为阻力与物体的速度成正比. 比例系数当然与物体的形状有关. 人们有时尽量把物体的形状做成流线型的 (炸弹), 而在另一些情况下 (降落伞) 却有恰恰相反的目的. 将作用在物体上的力加以比较, 我们得到下述方程, 它是物体在空气中降落的速度应满足的方程:

$$m\dot{v}(t)=mg-\alpha v. \tag{15}$$

以 m 除此方程并用 β 代表 $\frac{\alpha}{m}$, 最后得到

$$\dot{v}(t)=-\beta v+g. \tag{13'}$$

我们得到一个只是字母表示与 (13) 不同的方程. 我们指出, 若令 $-\beta v(t)+g=f(t)$, 则因 $f'(t)=-\beta\dot{v}(t)$, 那么从 (13′) 可得与它等价的方程

$$f'(t)=-\beta f(t). \tag{10′}$$

它与方程 (8) 或 (10) 是一样的, 只不过字母表示不同而已. 于是我们得到它的解

$$f(t)=f(0)e^{-\beta t}.$$

由此推出方程 (13′) 的解是

$$v(t)=\frac{1}{\beta}g+\left(v_0-\frac{1}{\beta}g\right)e^{-\beta t},$$

从而基本方程 (15) 的解是

$$v(t)=\frac{m}{\alpha}g+\left(v_0-\frac{m}{\alpha}g\right)e^{-\frac{\alpha}{m}t}, \tag{16}$$

其中 $v_0=v(0)$ 是物体的初始竖直速度.

从 (16) 看出, 当 $\alpha>0$ 时, 空气中的落体将逼近于具固定速率的运动, 且 $v(t)\approx\frac{m}{\alpha}g$. 这样看来, 与真空中的降落不同, 在空气中降落的速度不仅与物体的形状有关, 而且与它的质量也有关系. 当 $\alpha\to 0$ 时, 等式 (16) 的右端趋于 v_0+gt, 即趋于当 $\alpha=0$ 时从 (15) 得到的方程 (14) 的解.

利用公式 (16) 可以对空气中落体的极限速度的快慢有个概念.

例如, 如果期望降落伞能使中等体格的人以 10m/s 左右的速度开伞着地, 那么, 当持续自由下落速度达到大约 50m/s 时开伞, 跳伞者再过 3s 就将有大约 12m/s 的速度.

实际上, 从所给的条件和关系式 (16), 求得 $\frac{m}{\alpha}g\approx 10, \frac{m}{\alpha}\approx 1, v_0=50$, 因此关系式 (16) 具体化为

$$v(t)=10+40e^{-t}.$$

由于 $e^3\approx 20$, 所以当 $t=3$ 时得 $v(t)\approx 12(\mathrm{m/s})$.

5. 再谈数 e 及指数函数 $\exp x$ 我们已经通过例子证明 (也见于本节末尾的问题 3.4), 一系列自然现象从数学的观点上都用同一个微分方程

$$f'(x)=\alpha f(x) \tag{17}$$

来描述, 只要给出 “初始条件” $f(0)$, 这个方程的解就单值地确定了. 那时

$$f(x)=f(0)e^{\alpha x}.$$

我们曾经相当形式地引入了数 e 及函数 $e^x=\exp x$, 当时我们敷衍地说这实在是一个重要的数和重要的函数. 现在我们已经明白, 即使我们早先不曾引入这个函

数, 它无疑也一定要作为方程 (17) 的解而被引入. 方程 (17) 虽然简单得很, 但却是非常重要的. 确切地说, 只要对于某个具体的 α 值, 例如 $\alpha = 1$, 引入一个作为方程 (17) 的解的函数就够了. 这是因为, 一般的方程 (17) 经关系式 $x = \frac{t}{\alpha}(\alpha \neq 0)$ 而转换到新变元 t 后, 就归结到 $\alpha = 1$ 的情形.

实际上, 那时

$$f(x) = f\left(\frac{t}{\alpha}\right) = F(t), \frac{df(x)}{dx} = \frac{\frac{dF(t)}{dt}}{\frac{dx}{dt}} = \alpha F'(t).$$

于是代替方程 $f'(x) = \alpha f(x)$, 现在有 $\alpha F'(t) = \alpha F(t)$, 即

$$F'(t) = F(t).$$

于是, 我们来考察方程

$$f'(x) = f(x) \tag{18}$$

并用 $\exp x$ 来代表它的使 $f(0) = 1$ 的解.

我们来大略估量一下, 这个定义与先前关于函数 $\exp x$ 的定义是不是一致的.

我们尝试着从 $f(0) = 1$ 及 f 满足方程 (18) 这些条件出发来计算 $f(x)$ 的值. 由于 f 是可微函数, 所以 f 是连续的, 这样一来, 根据 (18) 函数 f' 也是连续的. 另外从 (18) 还推出 f 还有二阶导数 $f''(x) = f'(x)$, 进而一般地从 (18) 推出 f 是无穷可微函数. 由于函数 $f(x)$ 的变化速度 $f'(x)$ 是连续的, 所以在自变量变化的小间隔 h 中函数 f' 变化很小, 因此

$$f(x_0 + h) = f(x_0) + f'(\xi)h \approx f(x_0) + f'(x_0)h.$$

使用这个近似公式, 以小的步长 $h = \frac{x}{n}$ 来走过从 0 到 x 的线段, 其中 $n \in \mathbb{N}$. 如果 $x_0 = 0, x_{k+1} = x_k + h$. 那么, 我们将有

$$f(x_{k+1}) \approx f(x_k) + f'(x_k)h.$$

注意到 (18) 和条件 $f(0) = 1$, 有

$$\begin{aligned} f(x) = f(x_n) &\approx f(x_{n-1}) + f'(x_{n-1})h = f(x_{n-1})(1+h) \\ &\approx (f(x_{n-2}) + f'(x_{n-2})h)(1+h) = f(x_{n-2})(1+h)^2 \approx \cdots \\ &\approx f(x_0)(1+h)^n = f(0)(1+h)^n \\ &= \left(1 + \frac{x}{n}\right)^n. \end{aligned}$$

步长 $h = \frac{x}{n}$ 越小, 近似公式 $f(x) \approx \left(1 + \frac{x}{n}\right)^n$ 就越精确, 看起来这是一件十分自然

的事情 (确实也是可以证明的).

于是我们得到

$$f(x)=\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n.$$

特别地, 如果把量

$$f(1)=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n$$

记作 e 并证明 $e\neq 1$ 的话, 我们就得到

$$f(x)=\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n=\lim_{t\to 0}(1+t)^{\frac{x}{t}}=\lim_{t\to 0}\left[(1+t)^{\frac{1}{t}}\right]^x=e^x, \tag{19}$$

这是因为我们已经知道当 $u\to v$ 时 $u^\alpha\to v^\alpha$.

微分方程 (18) 的这种数值解法使我们求得公式 (19). 这种方法也是欧拉提出的, 叫作*欧拉折线法*. 这个名称与这个算法的几何意义相关: 用一条折线来代替方程的解 $f(x)$, 确切地说, 是近似地代替解 $f(x)$ 的图像, 这条折线在相应区段 $[x_k,x_{k+1}](k=0,1,\cdots,n-1)$ 上的那一节是由方程

$$y=f(x_k)+f'(x_k)(x-x_k)$$

给定的 (图46).

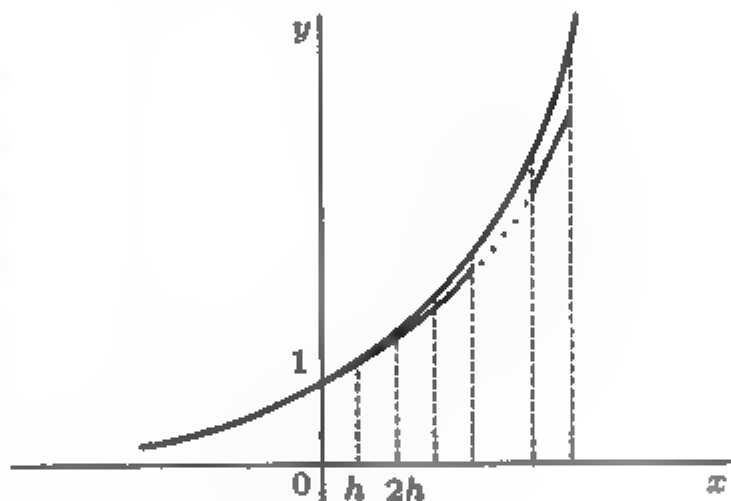


图 46

我们还遇到过函数 $\exp x$ 作为幂级数 $\sum\limits_{n=0}^{\infty}\frac{1}{n!}x^n$ 的和的定义. 也可以从方程 (18) 引出这个定义, 这只要使用下述常用的所谓*待定系数法*. 我们来求方程 (18) 的幂级数

$$f(x)=c_0+c_1x+\cdots+c_nx^n+\cdots \tag{20}$$

和的形式的解, 此幂级数的系数是待定的.

我们已经见到 (见 §5 定理 1), 从 (20) 推出 $c_n=\frac{f^{(n)}(0)}{n!}$. 但根据 (18), $f(0)=f'(0)=\cdots=f^{(n)}(0)=\cdots$, 并且, 由于 $f(0)=1$, 所以有 $c_n=\frac{1}{n!}$, 即如果解具有 (20) 的形式且 $f(0)=1$, 则必有

$$f(x)=1+\frac{1}{1!}x+\frac{1}{2!}x^2+\cdots+\frac{1}{n!}x^n+\cdots.$$

可以单独验证这个级数定义的函数确实可微 (不单在 $x=0$ 处), 而且它满足方程 (18) 和初始条件 $f(0)=1$. 不过我们不去做这件事了, 因为我们的目的只在于考察一下作为方程 (18) 在初始条件 $f(0)=1$ 下的解而引入的指数函数与早先我们所谈的指数函数 $\exp x$ 是否一致.

我们指出, 方程 (18) 还可以在复数域中来研究, 即认为 x 是任意的复数. 此时, 所做的全部论述依然有效, 只是可能部分地失去了欧拉方法的几何直观性.

于是, 自然也是我们期望的, 函数

$$e^z = 1 + \frac{1}{1!}z + \frac{1}{2!}z^2 + \cdots + \frac{1}{n!}z^n + \cdots$$

是方程

$$f'(z) = f(z)$$

的满足条件 $f(0) = 1$ 的唯一的解.

6. 振动 如果挂在弹簧下面的物体偏离了平衡位置, 例如先把它稍稍托起再放下, 那么, 它就要在平衡位置附近振动. 我们以一般的形式来描述这个过程.

设知道质量为 m 的质点可沿数轴 Ox 移动, 它被作用着与它和坐标原点的距离成正比的力 $F = -kx$ ①. 还知道质点的初始位置 $x_0 = x(0)$ 和它的初始速度 $v_0 = \dot{x}(0)$. 求质点的位置对时间的依赖关系 $x = x(t)$.

根据牛顿定律, 可把这个问题写成下述纯数学的形式: 求方程

$$m\ddot{x}(t) = -kx(t) \tag{21}$$

满足初始条件 $x(0) = x_0, \dot{x}(0) = v_0$ 的解.

把方程 (21) 改写成

$$\ddot{x}(t) + \frac{k}{m}x(t) = 0, \tag{22}$$

并且再用指数函数来试一试, 也就是说, 试着选一数 λ, 使函数 $x(t) = e^{\lambda t}$ 满足方程 (22).

把 $x(t) = e^{\lambda t}$ 代入 (22), 得到

$$\left(\lambda^2 + \frac{k}{m}\right)e^{\lambda t} = 0,$$

则

$$\lambda^2 + \frac{k}{m} = 0, \tag{23}$$

即 $\lambda_1 = -\sqrt{-\frac{k}{m}}, \lambda_2 = \sqrt{-\frac{k}{m}}$. 由于 $m > 0$, 所以当 $k > 0$ 时, 我们得到两个纯虚数 $\lambda_1 = -i\sqrt{\frac{k}{m}}, \lambda_2 = i\sqrt{\frac{k}{m}}$. 我们并没指望得到这样的结果. 虽说如此, 我们还是要继续作进一步的研究. 按照欧拉公式

$$e^{-i\sqrt{\frac{k}{m}}t} = \cos\sqrt{\frac{k}{m}}t - i\sin\sqrt{\frac{k}{m}}t,$$

$$e^{+i\sqrt{\frac{k}{m}}t} = \cos\sqrt{\frac{k}{m}}t + i\sin\sqrt{\frac{k}{m}}t.$$

①对于弹簧, 刻画它的弹性的系数 $k > 0$, 叫做弹簧的弹性系数.

因为关于实的时间 t 进行微分时, 函数 $e^{\lambda t}$ 的实部和虚部分别单独进行微分, 所以函数 $\cos\sqrt{\frac{k}{m}}t$ 及函数 $\sin\sqrt{\frac{k}{m}}t$ 应该单独地满足方程 (22). 实际上确是如此, 这容易直接验证. 这样, 复指数帮我们猜中了方程 (22) 的两个解. 这两个解的线性组合

$$x(t) = c_1\cos\sqrt{\frac{k}{m}}t + c_2\sin\sqrt{\frac{k}{m}}t \tag{24}$$

显然还是方程 (22) 的解.

我们从条件

$$\begin{aligned}
x_0 &= x(0) = c_1,\\
v_0 &= \dot{x}(0) = \left(-c_1\sqrt{\frac{k}{m}}\sin\sqrt{\frac{k}{m}}t + c_2\sqrt{\frac{k}{m}}\cos\sqrt{\frac{k}{m}}t\right)\Bigg|_{t=0}\\
&= c_2\sqrt{\frac{k}{m}}
\end{aligned}$$

来选取系数 c_1, c_2.

这样一来, 周期函数

$$x(t) = x_0\cos\sqrt{\frac{k}{m}}t + v_0\sqrt{\frac{m}{k}}\sin\sqrt{\frac{k}{m}}t \tag{25}$$

是所求的解.

式 (25) 经规格化的变换后可改写成

$$x(t) = \sqrt{x_0^2 + v_0^2\frac{m}{k}}\sin\left(\sqrt{\frac{k}{m}}t + \alpha\right), \tag{26}$$

式中

$$\alpha = \arcsin\frac{x_0}{\sqrt{x_0^2 + v_0^2\frac{m}{k}}}.$$

因此, 当 $k > 0$ 时, 质点将作周期振动, 周期是 $T = 2\pi\sqrt{\frac{m}{k}}$, 也就是说, 频率 $\frac{1}{T} = \frac{1}{2\pi}\sqrt{\frac{k}{m}}$, 而振幅是 $\sqrt{x_0^2 + v_0^2\frac{m}{k}}$. 我们确信这个结论, 因为, 从物理上考虑, 很明显, 所提问题的解 (25) 是唯一的.(见本节末尾的问题 5.)

函数 (26) 所描述的运动叫做**简谐振动**, 而方程式 (22) 叫做**简谐振动方程**.

现在回到方程 (23) 中 $k < 0$ 的情形. 此时, 函数

$$e^{\lambda_1 t} = e^{-\sqrt{\frac{-k}{m}}t} \text{和} e^{\lambda_2 t} = e^{\sqrt{\frac{-k}{m}}t}$$

二者都是方程 (22) 的实解, 且函数

$$x(t) = c_1e^{\lambda_1 t} + c_2e^{\lambda_2 t} \tag{27}$$

也是解. 常数 c_1、c_2 由下面的条件来决定:

$$\begin{cases} x_0 = x(0) = c_1 + c_2, \\ v_0 = \dot{x}(0) = c_1\lambda_1 + c_2\lambda_2. \end{cases}$$

所得的方程组总是单值可解的, 因为它的行列式 $\lambda_2 - \lambda_1 \neq 0$.

当 $k < 0$ 时, 因为数 λ_1 和 λ_2 反号, 所以从 (27) 看到, 如果 x_0 或 v_0 不为零的话, 力 $F = -kx$ 不仅不驱使质点回到平衡位置 $x = 0$, 而是使它随着时间的推移无限地远离平衡位置. 亦即, 在这种情况下, $x = 0$ 是不稳定平衡点.

最后, 我们来研究方程 (21) 的一个十分自然的改进, 通过这个改进更加鲜明地看出指数函数以及把基本初等函数联系起来的欧拉公式的好处.

假设我们所考察的质点是在介质 (在空气中或者在液体中) 运动, 介质的阻力不容忽略. 设阻力与质点的速度成正比. 那么代替方程 (21) 我们应该得到方程

$$m\ddot{x}(t) = -\alpha\dot{x}(t) - kx(t),$$

把它改写成

$$\ddot{x}(t) + \frac{\alpha}{m}\dot{x}(t) + \frac{k}{m}x(t) = 0. \tag{28}$$

如果还是求形如 $x(t) = e^{\lambda t}$ 的解, 那么我们得到二次方程

$$\lambda^2 + \frac{\alpha}{m}\lambda + \frac{k}{m} = 0,$$

它的根是

$$\lambda_{1、2} = -\frac{\alpha}{2m} \pm \sqrt{\frac{\alpha^2 - 4mk}{2m}}.$$

当 $\alpha^2 - 4mk > 0$ 时, 得到两个实根 λ_1、λ_2, 于是可求得形如 (27) 的解.

我们对 $\alpha^2 - 4mk < 0$ 的情况更感兴趣, 下边对它做更详细的考察. 这时, 根 λ_1、λ_2 都是复数 (不是实数也不是纯虚数):

$$\lambda_1 = -\frac{\alpha}{2m} - i\frac{\sqrt{4mk - \alpha^2}}{2m},$$
$$\lambda_2 = -\frac{\alpha}{2m} + i\frac{\sqrt{4mk - \alpha^2}}{2m}.$$

在这种情况下, 欧拉公式给出

$$e^{\lambda_1 t} = e^{-\frac{\alpha}{2m}t}(\cos\omega t - i\ \sin\omega t),$$
$$e^{\lambda_2 t} = e^{-\frac{\alpha}{2m}t}(\cos\omega t + i\ \sin\omega t),$$

其中 $\omega = \dfrac{\sqrt{4mk - \alpha^2}}{2m}$. 于是我们求得方程 (28) 的两个实解 $e^{-\frac{\alpha}{2m}t}\cos\omega t$ 和 $e^{-\frac{\alpha}{2m}t}$

$\sin\omega t$. 要想简单地猜出这两个解, 实在是太困难了. 现在来做它们的线性组合

$$x(t) = e^{-\frac{\alpha}{2m}t}(c_1\cos\omega t + c_2\sin\omega t), \tag{29}$$

选择满足初始条件 $x(0) = x_0$、$\dot{x}(0) = v_0$ 的 c_1、c_2, 从而求出原来问题的形如 (29) 的解.

这时, 可以验证, 所得的方程组总是单值可解的. 于是, 经过一些变换, 从 (29) 得到问题的如下形式的解:

$$x(t) = Ae^{-\frac{\alpha}{2m}t}\sin(\omega t + a), \tag{30}$$

式中 A 和 a 是由初始条件决定的常数.

从这个公式看到, 由于有因子 $e^{-\frac{\alpha}{2m}t}$, 其中 $\alpha > 0$、$m > 0$, 所以在所考察的情况下, 振动是衰减的, 叫阻尼振动, 其振幅衰减的速度依赖于比值 $\frac{\alpha}{m}$. 振动的频率

$$\frac{1}{2\pi}\omega = \frac{1}{2\pi}\sqrt{\frac{k}{m} - \left(\frac{\alpha}{2m}\right)^2}$$

不随时间变化. 量 ω 也只依赖于比值 $\frac{k}{m}$ 和 $\frac{\alpha}{m}$, 这一点其实早就可以从原来的方程 (28) 看出来. 当 $\alpha = 0$ 时, 我们重新回到无阻尼的简谐振动 (26) 和方程 (22).

练　习

1. 喷气运动的有效作用系数

a) 设 Q 是单位质量的火箭燃料的化学能, ω 是燃料喷出的速度. 那么 $\frac{1}{2}\omega^2$ 是喷出的单位质量燃料的动能. 在等式 $\frac{1}{2}\omega^2 = \alpha Q$ 中的系数 α 是燃料燃烧和喷出过程的有效作用系数. 对于固体燃料 (无烟火药) $\omega = 2\text{km/s}$, $Q = 1000\text{kcal/kg}$, 而对于液体燃料 (带氧汽油) $\omega = 3\text{km/s}$, $Q = 2500\text{kcal/kg}$. 在这些情况下确定系数 α.

b) 火箭的有效作用系数用它的终极动能 $m_k \cdot \frac{v^2}{2}$ 与烧掉的燃料的化学能 $m_T Q$ 的比来确定. 使用公式 (4), 求出火箭的有效作用系数用 m_K、m_T、Q 和 α(见 a)) 表示的公式.

c) 如果带有液体喷气发动机的汽车加速到规定的 60km/h 的市区行车速度, 估算它的有效作用系数.

d) 估算把卫星送入靠近地面的低轨道的液体燃料火箭的有效作用系数.

e) 估算对于怎样的终极速度, 液体燃料喷气运动有最大的有效作用系数.

f) 指出, 要使不管具有何种形式的燃料的火箭得到尽可能大的有效作用系数, 燃料和火箭本体的质量比 $\frac{m_T}{m_K}$ 应该取怎样的值.

2. 气压公式

a) 用本节 2 段所得结果, 求出把空气柱的温度对压强的影响考虑进去的修正项的公式, 设空气的温度经常变化着 (例如季节性的变化), 变化的范围是 ±40 °C.

b) 按公式 (9) 求出当温度为 −40 °C、0 °C、40 °C 时压强对高度的依赖关系, 并将这些结果与您在 a) 中所得的近似公式所给出的结果加以比较.

c) 设柱中空气的温度随高度按规律 $T'(h) = -\alpha T_0$ 变化, 其中 T_0 是地面上的空气温度, 而 $\alpha \approx 7 \times 10^{-7} \frac{1}{\text{cm}}$. 由这些条件导出压强对高度的依赖关系.

d) 按公式 (9), 也按您在 c) 中所得的公式, 分别求出在深 1km、3km、9km 的矿井中的压强.

e) 不管高度如何, 空气中大约含着 $\frac{1}{5}$ 的氧气. 氧气的部分压强也大约占空气压强的 $\frac{1}{5}$. 某种鱼可以在氧气分压不低于 0.15 个大气压的条件下生活. 能不能指望在与海洋平面同高度的河中见到这种鱼? 能不能在流人高度为 3.81km 的的的喀喀湖[①] 的小河中见到这种鱼?

3. 放射衰变

a) 在地球的矿石中取样测量放射性物质的含量以及它衰变的生成物的含量, 并总是认为初始时不曾有衰变生成物, 这样就可以大约估算出地球的年龄 (不管在什么情况下, 总是从所说的放射性物质产生的时刻算起). 设在矿石中放射性物质的质量是 m_r, 它衰变的生成物质量是 r_r. 知道物质的半衰期 T, 求从衰变开始以来经历的时间以及在初始时刻样品中放射性物质的量.

b) 在矿石中镭原子约占一切原子总量的 10^{-12}. 在 10^5、10^9 和 5×10^9 年之前镭的含量如何?(一般认为地球的年龄是 5×10^9 年).

c) 在诊断肾脏疾病时, 常常把一些特殊物质, 例如肌酸引人有机体内, 然后测定肾脏把这些物质从血液中排除出去的能力 ("廓清试验"). 与这类过程相反的例子是恢复供血者或突然大出血的病人的血红素的浓度. 在所有这些情况下, 引人物质排出的量 (或者, 反过来, 不足的数量的弥补) 遵从于规律 $N = N_0 e^{-\frac{t}{\tau}}$, 其中 N 是初始引人量为 N_0 的物质经过时间 t 后在有机体内尚存的量 (或者说是分子数目); 而 τ 是所谓时间常数, 它表示这样一段时间, 经过这段时间后, 有机体内尚存引人物的量为原始引人量 N_0 的 $\frac{1}{e}$. 容易验证, 时间常数是半存留时间 (或半衰期) 的 1.44 倍, 所谓半存留时间是有机体内尚存引人物成为原始引人量的 $\frac{1}{2}$ 时所历经的时间.

设放射性物质从有机体内排出的速度由时间常数 τ_0 来表征, 同时它还以时间常数 τ_p 进行着衰变. 证明在这种情况下, 刻画物质在体内存留时间长短的时间常数 τ 由关系式 $\tau^{-1} = \tau_0^{-1} + \tau_p^{-1}$ 来确定.

d) 从供血者取了某一数量的血液, 这些血中含有 201mg 铁; 为了弥补失去的铁, 让供血者服用硫酸铁剂药片, 每天三次连服七天. 每片药含铁 67mg. 供血者血中铁的含量按指数律恢复到正常, 指数律的时间常数大约是七昼夜. 设在取血之后, 药片中的铁以最大速度一下子就进人血中. 请确定, 在恢复到血液正常含铁量的全部时间中, 药片中所含的铁大约有多大一部分进人血液中.

e) 为了诊断的目的, 给恶性肿瘤患者服一定量的放射性磷 P^{32}, 然后每过一段相同的时间, 测量一下大腿皮肤处的放射性强度. 放射性强度的减弱遵从指数规律. 因为已经知道磷的半衰期是 14.3 昼夜, 所以按照测量得到的资料, 就能确定单由生物方面的因素引起的放射性强度减弱过程的时间常数. 如果通过观测已确定整体上的放射性强度减弱过程的

①的的喀喀湖属秘鲁和玻利维亚. —— 译者注.

时间常数是 9 4 昼夜, 求由生物原因单独引起的放射性强度减弱的时间常数 (见 c)).

4. **辐射的吸收** 射线穿过介质时总要被介质吸收掉一部分. 在很多场合 (线性理论) 可以认为, 通过厚度为两个单位的介质层和接连通过两个厚度为一个单位的介质层, 辐射的减弱是一样的

a) 证明: 在所说的条件下, 辐射的吸收遵从规律 $I = I_0 e^{-kl}$, 其中 I_0 是落到吸收物质上的辐射强度, I 是通过厚 l 的吸收物质层之后的辐射强度, 而 k 是系数, 它的量纲是波长量纲的倒数.

b) 在光线被水吸收的情况, 系数 k 依赖于射到水中的光线的波长, 例如: 紫外线, $k = 1.4 \times 10^{-2}$; 蓝光, $k = 4.6 \times 10^{-4}$; 绿光, $k = 4.4 \times 10^{-4}$; 红光, $k = 2.9 \times 10^{-3}$. 太阳光垂直地照在深 10m 的洁净的湖面上. 把太阳光的上列每种成分在湖面上的强度与在湖底处的强度进行比较.

5. 证明: 如果点的运动规律 $x = x(t)$ 满足简谐振动方程 $m\ddot{x} + kx = 0$, 那么

a) 量 $E = \frac{m\dot{x}^2(t)}{2} + \frac{kx^2(t)}{2}$ 是常数 $\left(E = K + U \text{是点在时刻} t \text{的动能} K = \frac{m\dot{x}^2(t)}{2} \text{与势能} U = \frac{kx^2(t)}{2} \text{的和}\right)$.

b) 如果 $x(0) = 0$ 且 $\dot{x}(0) = 0$, 那么 $x(t) \equiv 0$.

c) 存在唯一的一个满足初始条件 $x(0) = x_0$ 且 $\dot{x}(0) = v_0$ 的运动

$$x = x(t)$$

d) 验证: 如果点在有摩擦的介质中运动且 $x = x(t)$ 满足方程

$$m\ddot{x} + \alpha\dot{x} + kx = 0,$$

$\alpha > 0$, 那么, 量 E(见 a)) 是递减的. 求出 E 递减的速度, 并根据 E 的物理意义解释所得结果的物理意义.

6. 在胡克[①] 中心力作用下的运动 (*平面振子*).

作为第 6 段和习题 6 中线性振子方程 (21) 的发展, 我们将考察方程 $m\ddot{\boldsymbol{r}}(t) = -k\boldsymbol{r}(t)$, 它是在中心力作用下质量为 m 的质点的向径 $\boldsymbol{r}(t)$ 所满足的方程, 这个中心力与质点到中心点的距离 $|\boldsymbol{r}(t)|$ 成正比, 比例系数为 $k > 0$. 如果质点与中心是胡克弹性联结, 譬如, 用一个弹性系数为 k 的弹簧联结, 就产生这种力.

a) 微分向量积 $\boldsymbol{r}(t) \times \dot{\boldsymbol{r}}(t)$, 将证明, 质点是在过中心且包含向量 $\boldsymbol{r}_0 = \boldsymbol{r}(t_0)$, $\dot{\boldsymbol{r}}_0 = \dot{\boldsymbol{r}}(t_0)$, 即质点的初始位置向量和初始速度向量的平面内运动 (平面振子) . 如果向量 $\boldsymbol{r}_0 = \boldsymbol{r}(t_0), \dot{\boldsymbol{r}}_0 = \dot{\boldsymbol{r}}(t_0)$ 共线, 则运动在包含中心和向量 $\boldsymbol{r}_0$ 的直线内进行 (第 6 段中研究的线性振子).

b) 验证: 平面振子的轨道是椭圆, 而且, 质点沿它作周期运动. 试求回转周期.

c) 试证: 量 $E = m\dot{\boldsymbol{r}}^2(t) + k\boldsymbol{r}^2(t)$ 守恒 (不随时间变化).

d) 试证: 初始条件 $\boldsymbol{r}_0 = \boldsymbol{r}(t_0), \dot{\boldsymbol{r}}_0 = \dot{\boldsymbol{r}}(t_0)$ 完全确定质点在随后时间的运动.

---

①胡克 (Hooke, RoBert)(1635—1703) —— 英国自然科学家, 多方面的学者和实验物理学家. 他发现了组织的细胞结构,"细胞" 的术语也是他引人的. 他是弹性的数学理论和光的波动理论的创始人之一, 他提出了引力假定和引力相互作用的反平方定律.

7. 行星轨道的椭圆性.

上题说明, 在中心胡克力作用下的质点运动是在平面内进行的运动. 设这是复变量 $z = x + iy$ 的平面, 运动由两个实函数 $x = x(t), y = y(t)$, 或等价地, 由一个时间 t 的复函数 $z = z(t)$ 定义. 为了简单, 在习题 6 中设 $m = 1, k \approx 1$, 考察最简单的这种运动的方程 $\ddot{z}(t) = -z(t)$.

a) 在从习题 6 知道了这个方程满足初始条件 $z_0 = z(t_0), \dot{z}_0 = \dot{z}(t_0)$ 的解是唯一的, 试求它的形如 $z(t) = c_1 e^{it} + c_2 e^{-it}$ 的解, 并利用欧拉公式再次验证, 运动的轨道是以零为中心的随圆.(在一些特定情况它变成圆周或退化成闭区间 —— 试分析是哪些情况.)

b) 注意到量 $|\dot{z}(t)|^2 + |z(t)|^2$ 在点 $z(t)$ 的运动过程中不变, 这里 $z(t)$ 满足方程 $\ddot{z}(t) = -z(t)$, 试验证, 点 $w(t) = z^2(t)$ 关于新参数 (时间) τ 满足方程 $\frac{d^2 w}{d\tau^2} = -c\frac{w}{|w|^3}$, 这里 c 是常数, $w = w(t(\tau))$, 而 τ 和 t 由 $\tau = \tau(t)$(或 $t = t(\tau)$) 联系. 其中 $\tau(t)$ 满足方程 $\frac{d\tau}{dt} = |z(t)|^2$. 这样一来, 胡克中心力场中的运动与牛顿引力场中的运动原来是密切相关的.

c) 把这里的结果与 §5 的习题 8 的结果进行比较并证明行星轨道的椭圆性.

d) 如果你会用计算机, 那么, 再读一遍在第 5 段中叙述的欧拉折线法, 关于原点, 用这个方法计算几个 e^x 的值.(注意, 除微分定义外, 更准确地说, 除公式 $f(x_n) \approx f(x_{n-1}) + f'(x_{n-1})h (h = x_n - x_{n-1})$ 外, 这个方法没有用到什么.

现设 $\boldsymbol{r}(t) = (x(t), y(t)), \boldsymbol{r}_0 = \boldsymbol{r}(0) = (1,0), \dot{\boldsymbol{r}}_0 = \dot{\boldsymbol{r}}(0) = (0,1)$ 且 $\ddot{\boldsymbol{r}}(t) = -\boldsymbol{r}(t)/|\boldsymbol{r}(t)|^3$. 借助公式

$$\boldsymbol{r}(t_n) \approx \boldsymbol{r}(t_{n-1}) + \boldsymbol{v}(t_{n-1})h,$$
$$\boldsymbol{v}(t_n) \approx \boldsymbol{v}(t_{n-1}) + \boldsymbol{a}(t_{n-1})h,$$

这里 $\boldsymbol{v}(t) = \dot{\boldsymbol{r}}(t), \boldsymbol{a}(t) = \dot{\boldsymbol{v}}(t) = \ddot{\boldsymbol{r}}(t)$, 试用欧拉公式计算点的运动轨道, 研究它有怎样的形状, 它是怎样由点随着时间的推移形成的.

§7. 原函数

在微分学中, 要学会求函数的微分以及写出函数的导数之间的关系, 与此同时, 学会从函数的导数所满足的关系式求出原来的函数, 也同样是很有价值的. 这一点我们在上节的例子中已经认识到了. 这类问题中最简单的, 然而以后会见到, 也是极重要的, 是根据一个函数 $F(x)$ 的已知导数 $F'(x) = f(x)$ 来求函数 $F(x)$ 的问题. 本节就来对这个问题做些初步的讨论.

1. 原函数和不定积分

定义 1 *函数 $F(x)$ 叫做函数 $f(x)$ 在某个区间上的原函数, 如果在这个区间上函数 $F(x)$ 可微且满足方程 $F'(x) = f(x)$, 也就是说满足关系式 $dF(x) = f(x)dx$.*

例 1 函数 $F(x) = \operatorname{arctg} x$ 是 $f(x) = \frac{1}{1+x^2}$ 在全数轴上的原函数, 因为 $\operatorname{arctg}' x = \frac{1}{1+x^2}$.

例 2 函数 $F(x) = \operatorname{arcctg} \frac{1}{x}$ 既是函数 $f(x) = \frac{1}{1+x^2}$ 在正半轴上的原函数，也是它在负半轴上的原函数，因为当 $x \neq 0$ 时，

$$F'(x) = -\frac{1}{1+\left(\frac{1}{x}\right)^2}\left(-\frac{1}{x^2}\right) = \frac{1}{1+x^2} = f(x).$$

原函数是否存在？已知函数的原函数的全体构成怎样的一个集合？

在积分学中将证明这样一个基本事实，即任何在区间上连续的函数在此区间上都有原函数.

我们只是告诉读者有这样一个事实，而在这节所用到的，本质上只是下述我们已知的 (见第五章 §3 第 1 段) 关于已给函数在实数区间上的原函数的集合的性质，这种性质是从拉格朗日定理得到的.

命题 1 *如果 $F_1(x)$ 和 $F_2(x)$ 都是函数 $f(x)$ 在同一个区间上的原函数，那么，它们的差 $F_1(x) - F_2(x)$ 在这个区间上是常数.*

$F_1(x)$ 和 $F_2(x)$ 的比较必须在连通区间上进行，这一点，如在证明这一命题时所指出的那样 (见第五章 §3 第 1 段)，是很重要的. 这也可以从例 1 和例 2 的对比中看出来，它们所讨论的两个函数

$$F_1(x) = \operatorname{arctg} x \text{ 和 } F_2(x) = \operatorname{arcctg} \frac{1}{x}$$

的导数在同时有定义的区域 $\mathbb{R}\backslash 0$ 上重合，但是

$$F_1(x) - F_2(x) = \operatorname{arctg} x - \operatorname{arcctg} \frac{1}{x}$$

当 $x > 0$ 时恒为零，而当 $x < 0$ 时恒为 $-\pi$，这是因为当 $x > 0$ 时，$\operatorname{arcctg} \frac{1}{x} = \operatorname{arctg} x$，而当 $x < 0$ 时，$\operatorname{arcctg} \frac{1}{x} = \pi + \operatorname{arctg} x$.

求微分的运算有自己的名称 "微分法" 和自己的数学记号 $dF(x) = F'(x)dx$. 与此一样，求原函数的运算也有自己的名称 "不定积分法" 和自己的数学记号

$$\int f(x)dx. \tag{1}$$

称它为函数 $f(x)$ 在给定区间上的不定积分.

于是我们把记号 (1) 看作是表示函数 f 在所考虑的区间上的任何一个原函数的记号.

在记号 (1) 中，符号 $\int$ 叫做*不定积分号*，f 叫做*被积函数*，而 $f(x)dx$ 叫做*被积表达式*.

从命题 1 推出, 如果 $F(x)$ 是函数 $f(x)$ 在区间上的某一个具体的原函数, 那么在这个区间上

$$\int f(x)dx = F(x) + C, \tag{2}$$

即任何其他的原函数都可以从这个具体的 $F(x)$ 添加某个常数而得到.

如果 $F'(x) = f(x)$, 即 F 是 f 在某个区间上的原函数, 那么从 (2) 有

$$d\int f(x)dx = dF(x) = F'(x)dx = f(x)dx. \tag{3}$$

此外, 按照不定积分是原函数中的任意一个这样一个概念, 从 (2) 还推出

$$\int dF(x) = \int F'(x)dx = F(x) + C. \tag{4}$$

公式 (3) 和 (4) 建立了微分法和不定积分法这两种运算之间的关系. 这两个运算互为逆运算, 精确到出现在公式 (4) 中的不定常数 C.

至此, 我们只讨论了公式 (2) 中的常数 C 的数学属性. 现在我们用最简单的例子来指出它的物理意义. 设点沿直线运动, 它的速度 $v(t)$ 作为时间的函数是已知的 (例如 $v(t) \equiv v$). 如果 $x(t)$ 是点在时刻 t 的坐标, 那么函数 $x(t)$ 满足方程 $\dot{x}(t) = v(t)$, 即它是 $v(t)$ 的原函数. 能否根据某一段时间内的速度 $v(t)$ 就确定点在轴上的位置? 显然不能. 能够根据速度和时间来确定的, 是在这段时间内所通过的路程 s, 却不能确定在轴上的位置. 但是, 一旦指出了点在某一时刻, 例如 $t = 0$ 时的位置, 也就是说给定了初始条件 $x(0) = x_0$, 那么, 此点的位置就完全由速度和时间决定. 在没有给定初始条件时, 运动规律 $x(t)$ 可以是形如 $x(t) = \tilde{x}(t) + c$ 的运动规律中的任何一个, 其中 $\tilde{x}(t)$ 是函数 $v(t)$ 的任意一个具体的原函数, 而 c 是任意的常数. 但是在给定初始条件 $x(0) = x_0$ 之后, 整个不确定性就完全消失了, 因为我们应该有 $x(0) = \tilde{x}(0) + c = x_0$, 即 $c = x_0 - \tilde{x}(0)$, 从而 $x(t) = x_0 + [\tilde{x}(t) - \tilde{x}(0)]$. 这个公式完全是物理的, 因为在这个公式中, 任意的原函数 $\tilde{x}$ 须以差的形式出现, 这个差确定了从已知的初始位置 $x(0) = x_0$ 开始走过的路程或位移.

2. 求原函数的基本的一般方法 按定义, 不定积分的记号 (1) 代表导数等于被积函数的函数, 由此定义出发, 考虑到关系式 (2) 和微分法的定律, 可以断定下列关系式成立:

a.
$$\int (\alpha u(x) + \beta v(x))\, dx = \alpha \int u(x)dx + \beta \int v(x)dx + c. \tag{5}$$

b.
$$\int (uv)'(x)dx = \int u'(x)v(x)dx + \int u(x)v'(x)dx + c. \tag{6}$$

c. 若在某区间 I_x 上

$$\int f(x)dx = F(x) + c$$

而 $\varphi: I_t \to I_x$ 是区间 I_t 到 I_x 的光滑 (即连续可微) 映射, 那么

$$\int (f \circ \varphi)(t)\varphi'(t)dt = F \circ \varphi(t) + c. \tag{7}$$

等式 (5)、(6)、(7) 可经直接对左边和右边求微分来验证. 其间, 在 (5) 中用到微分运算的线性性质, 在 (6) 中用到乘积的微分法则, 而在 (7) 中使用复合函数的微分法则.

我们知道, 微分法则使我们可以对已知函数的线性组合、乘积以及复合来求微分. 与此类似, 在一系列场合, 关系式 (5)、(6)、(7) 使我们可以把求一些函数的原函数的运算归结为求更简单些的函数的原函数, 或导致原函数本来就是已知的问题. 有一批这样的已知的原函数, 可以编成如下简单的不定积分表. 这个表是由基本初等函数的导数表改写得来的 (见 §2 第 3 段):

$$\int x^\alpha dx = \frac{1}{\alpha+1}x^{\alpha+1} + c \quad (\alpha \neq -1),$$

$$\int \frac{1}{x}dx = \ln|x| + c,$$

$$\int a^x dx = \frac{1}{\ln a}a^x + c \quad (0 < a \neq 1),$$

$$\int e^x dx = e^x + c,$$

$$\int \sin x dx = -\cos x + c,$$

$$\int \cos x dx = \sin x + c,$$

$$\int \frac{1}{\cos^2 x}dx = \operatorname{tg} x + c,$$

$$\int \frac{1}{\sin^2 x}dx = -\operatorname{ctg} x + c,$$

$$\int \frac{1}{\sqrt{1-x^2}}dx = \begin{cases} \arcsin x + c, \\ -\arccos x + \tilde{c}, \end{cases}$$

$$\int \frac{1}{1+x^2}dx = \begin{cases} \operatorname{arctg} x + c, \\ -\operatorname{arcctg} x + \tilde{c}, \end{cases}$$

$$\int \operatorname{sh} x dx = \operatorname{ch} x + c,$$

$$\int \operatorname{ch} x dx = \operatorname{sh} x + c,$$

$$\int \frac{1}{\operatorname{ch}^2 x}dx = \operatorname{th} x + c,$$

$$\int \frac{1}{\operatorname{sh}^2 x}dx = -\operatorname{cth} x + c,$$

$$\int \frac{1}{\sqrt{x^2 \pm 1}} dx = \ln|x + \sqrt{x^2 \pm 1}| + c,$$
$$\int \frac{1}{1 - x^2} dx = \frac{1}{2} \ln\left|\frac{1 + x}{1 - x}\right| + c.$$

这些公式中的每个式子都是在实轴 $\mathbb{R}$ 上相应的被积函数定义的区间上来考虑的. 如果这样的区间有好几个, 那么式子右端的常数 c 可以在不同的区间上取不同的值.

现在来考察几个使用关系式 (5)、(6)、(7) 的例子.

先做出下述一般性的注释.

因为只要求出了定义在区间上的函数的某一个原函数, 其他原函数就可由添加常数得到, 所以, 为了简化写法, 下面各处只对作为已知函数的具体原函数所得到的最后结果添加任意常数.

a. 不定积分的线性性质 这个标题的意思应是, 根据关系式 (5), 函数的线性组合的原函数是这些函数的原函数的线性组合.

例 3

$$\begin{aligned}
&\int (a_0 + a_1 x + \cdots + a_n x^n) dx \\
&= a_0 \int 1 dx + a_1 \int x dx + \cdots + a_n \int x^n dx \\
&= c + a_0 x + \frac{1}{2} a_1 x^2 + \cdots + \frac{1}{n+1} a_n x^{n+1}.
\end{aligned}$$

例 4

$$\begin{aligned}
\int \left(x + \frac{1}{\sqrt{x}}\right)^2 dx &= \int \left(x^2 + 2\sqrt{x} + \frac{1}{x}\right) dx \\
&= \int x^2 dx + 2 \int x^{\frac{1}{2}} dx + \int \frac{1}{x} dx \\
&= \frac{1}{3} x^3 + \frac{4}{3} x^{\frac{3}{2}} + \ln|x| + c.
\end{aligned}$$

例 5

$$\begin{aligned}
\int \cos^2 \frac{x}{2} dx &= \int \frac{1}{2}(1 + \cos x) dx \\
&= \frac{1}{2} \int (1 + \cos x) dx = \frac{1}{2} \int 1 dx + \frac{1}{2} \int \cos x dx \\
&= \frac{1}{2} x + \frac{1}{2} \sin x + c.
\end{aligned}$$

b. 分部积分法 公式 (6) 可以改写成

$$u(x)v(x) = \int u(x) dv(x) + \int v(x) du(x) + c$$

也就是

$$\int u(x)dv(x) = u(x)v(x) - \int v(x)du(x) + c. \tag{6'}$$

这表明, 在求函数 $u(x)v'(x)$ 的原函数时, 问题可以归结为求函数 $u'(x)v(x)$ 的原函数, 把微分运算转到另一个因子上部分地对函数求积分, 且如 (6′) 表明的那样, 分出一项 $u(x)v(x)$. 公式 (6′) 叫做分部积分公式.

例 6

$$\begin{aligned}\int \ln x dx &= x\ln x - \int x d\ln x \\ &= x\ln x - \int x\frac{1}{x}dx = x\ln x - \int 1dx = x\ln x - x + c.\end{aligned}$$

例 7

$$\begin{aligned}\int x^2e^x dx &= \int x^2 de^x = x^2e^x - \int e^x dx^2 \\ &= x^2e^x - 2\int xe^x dx = x^2e^x - 2\int x de^x = x^2e^x - 2\left(xe^x - \int e^x dx\right) \\ &= x^2e^x - 2xe^x + 2e^x + c = (x^2 - 2x + 2)e^x + c.\end{aligned}$$

c. 不定积分中的变量替换 公式 (7) 表明在求函数 $f\circ\varphi(t)\varphi'(t)$ 的原函数时可如下处理:

$$\begin{aligned}\int f\circ\varphi(t)\varphi'(t)dt &= \int f(\varphi(t))d\varphi(t) = \int f(x)dx \\ &= F(x) + c = F(\varphi(t)) + c,\end{aligned}$$

也就是说, 先在积分号下进行替换 $\varphi(t) = x$ 而转到新变元 x, 然后作为 x 的函数求原函数, 再经过替换 $x = \varphi(t)$ 回到原来的变元 t.

例 8

$$\int \frac{tdt}{1+t^2} = \frac{1}{2}\int \frac{d(t^2+1)}{1+t^2} = \frac{1}{2}\int \frac{dx}{x} = \frac{1}{2}\ln|x| + c = \frac{1}{2}\ln(t^2+1) + c.$$

例 9

$$\begin{aligned}\int \frac{dx}{\sin x} &= \int \frac{dx}{2\sin\frac{x}{2}\cos\frac{x}{2}} = \int \frac{d\left(\frac{x}{2}\right)}{\operatorname{tg}\frac{x}{2}\cdot\cos^2\frac{x}{2}} \\ &= \int \frac{du}{\operatorname{tg}u\cos^2 u} = \int \frac{d\operatorname{tg}u}{\operatorname{tg}u} = \int \frac{dv}{v} = \ln|v| + c \\ &= \ln|\operatorname{tg}u| + c = \ln\left|\operatorname{tg}\frac{x}{2}\right| + c.\end{aligned}$$

我们考察了几个单独使用不定积分的性质 a、b、c 的例子. 事实上在大多数场合这些性质是同时使用的.

例 10

$$\begin{aligned}\int \sin 2x \cdot \cos 3x dx &= \frac{1}{2}\int (\sin 5x - \sin x)dx \\ &= \frac{1}{2}\left(\int \sin 5x dx - \int \sin x dx\right) = \frac{1}{2}\left(\frac{1}{5}\int \sin 5x d5x + \cos x\right) \\ &= \frac{1}{10}\int \sin u du + \frac{1}{2}\cos x = -\frac{1}{10}\cos u + \frac{1}{2}\cos x + c \\ &= \frac{1}{2}\cos x - \frac{1}{10}\cos 5x + c.\end{aligned}$$

例 11

$$\begin{aligned}\int \arcsin x dx &= x \arcsin x - \int x d \arcsin x \\ &= x\arcsin x - \int \frac{x}{\sqrt{1-x^2}}dx = x\arcsin x + \frac{1}{2}\int \frac{d(1-x^2)}{\sqrt{1-x^2}} \\ &= x\arcsin x + \frac{1}{2}\int u^{-\frac{1}{2}}du = x\arcsin x + u^{\frac{1}{2}} + c \\ &= x\arcsin x + \sqrt{1-x^2} + c.\end{aligned}$$

例 12

$$\begin{aligned}\int e^{ax}\cos bx dx &= \frac{1}{a}\int \cos bx de^{ax} \\ &= \frac{1}{a}e^{ax}\cos bx - \frac{1}{a}\int e^{ax} d\cos bx \\ &= \frac{1}{a}e^{ax}\cos bx + \frac{b}{a}\int e^{ax}\sin bx dx \\ &= \frac{1}{a}e^{ax}\cos bx + \frac{b}{a^2}\int \sin bx de^{ax} \\ &= \frac{1}{a}e^{ax}\cos bx + \frac{b}{a^2}e^{ax}\sin bx - \frac{b}{a^2}\int e^{ax}d\sin bx \\ &= \frac{1}{a}e^{ax}\cos bx + \frac{b}{a^2}e^{ax}\sin bx - \frac{b^2}{a^2}\int e^{ax}\cos bx dx.\end{aligned}$$

从得到的等式推出

$$\int e^{ax}\cos bx dx = \frac{a\cos bx + b\sin bx}{a^2+b^2}e^{ax} + c.$$

这个结果也可以使用欧拉公式以及函数 $e^{(a+ib)x} = e^{ax}\cos bx + ie^{ax}\sin bx$ 的原

函数是

$$\frac{1}{a+ib}e^{(a+ib)x}=\frac{a-ib}{a^2+b^2}e^{(a+ib)x}$$
$$=\frac{a\cos bx+b\sin bx}{a^2+b^2}e^{ax}+i\frac{a\sin bx-b\cos bx}{a^2+b^2}e^{ax}$$

这一事实得到.

留心此事, 将来有用. 对于实的 x, 上式容易直接验证, 而现在, 只要对函数 $\frac{1}{a+ib}e^{(a+ib)x}$ 的实、虚部分别求导数.

特别地, 由此还得到

$$\int e^{ax}\sin bx dx=\frac{a\sin bx-b\cos bx}{a^2+b^2}e^{ax}+c.$$

所考察的函数不多的各种例子表明, 即使在求初等函数的原函数时, 也常常要用到附加的变换和技巧. 这些变换和技巧, 在求那些我们已知其导数的函数的复合函数的导数时, 是全然未曾有过的. 发生这样的困难并不是偶然的. 例如, 与求导数不同, 对于初等函数求原函数的结果, 可能得到的已不再是初等函数的复合. 因此, 不应该把 "求原函数" 与 "用初等函数表出给定的初等函数的原函数" 混为一谈, 后者有时是不能实现的. 一般而言, 初等函数的范围是很窄的, 还有很多对于应用很重要的特殊函数. 我们说, 人们对这些特殊函数的研究及表格制作, 一点也不比 $\sin x$ 和 e^x 来得逊色.

例如,*积分正弦* $\mathrm{Si}x$ 是函数 $\frac{\sin x}{x}$ 的一个原函数 $\int\frac{\sin x}{x}dx$, 当 $x\to 0$ 时它趋于零. 这样的原函数是存在的, 但它与函数 $\frac{\sin x}{x}$ 的其他任何原函数一样, 都不是初等函数的复合.

类似地, 当 $x\to\infty$ 时趋于零的特定的函数

$$\mathrm{Ci}\,x=\int\frac{\cos x}{x}dx$$

不是初等的, 它叫做*积分余弦*.

函数 $\frac{1}{\ln x}$ 的原函数 $\int\frac{dx}{\ln x}$ 也是非初等的. 其中的一个用记号 $\mathrm{li}x$ 表示, 叫做*积分对数*. 它满足条件: 当 $x\to+0$ 时 $\mathrm{li}x\to 0$ (第六章 §5 中将更详细地谈及特殊函数 $\mathrm{Si}x$,$\mathrm{Ci}x$,$\mathrm{li}x$).

考虑到求原函数的这些困难, 人们编制了相当庞大的不定积分表. 不过为了顺利地使用这些表格, 也为了能在问题十分简单时避免查表, 我们必须掌握处理不定积分的某些技巧.

本节后面的部分研究几类特别的函数的求积问题, 这些函数的原函数都能表成初等函数的复合.

3. 有理函数的原函数 我们考虑形如 $\int R(x)dx$ 的积分的问题, 其中 $R(x) = \frac{P(x)}{Q(x)}$ 是多项式的比.

如果在实数域中考虑, 则从代数学中得知 (见 §5, 4 段的公式 (37)), 任何这样的分式都可展开成和式

$$\frac{P(x)}{Q(x)} = p(x) + \sum_{j=1}^{l}\left(\sum_{k=1}^{k_j}\frac{a_{jk}}{(x-x_j)^k}\right) + \sum_{j=1}^{n}\left(\sum_{k=0}^{m_j}\frac{b_{jk}x+c_{jk}}{(x^2+p_jx+q_j)^k}\right), \tag{8}$$

其中 $p(x)$ 是多项式 (它仅当 $P(x)$ 的次数不小于 $Q(x)$ 的次数时出现), a_{jk}、b_{jk}、c_{jk} 是单值确定的实数, 而 $Q(x) = (x-x_1)^{k_1} \cdot \cdots \cdot (x-x_l)^{k_l}(x^2+p_1x+q_1)^{m_1} \cdot \cdots \cdot (x^2+p_nx+q_n)^{m_n}$.

关于如何构作展开式 (8), 我们在 §5 已谈过了. 作出展开式 (8) 之后, 函数 $R(x)$ 的求积归结为单个被加项的求积.

在例 1 中我们已经求过多项式的积分, 因此剩下的只是考虑形如

$$\frac{1}{(x-a)^k} \text{和} \frac{bx+c}{(x^2+px+q)^k} \qquad (\text{其中} k \in \mathbf{N})$$

的分式的求积.

关于第一个分式, 问题立即得到解决, 因为

$$\int (x-a)^{-k}dx = \begin{cases} \frac{1}{-k+1}(x-a)^{-k+1} + c, & \text{当} k \neq 1, \\ \ln|x-a| + c, & \text{当} k = 1. \end{cases} \tag{9}$$

对于积分

$$\int \frac{bx+c}{(x^2+px+q)^k}dx$$

我们做如下的处理. 表多项式 x^2+px+q 成 $\left(x+\frac{p}{2}\right)^2 + \left(q-\frac{p^2}{4}\right)$, 其中 $q-\frac{p^2}{4} > 0$, 因为多项式 x^2+px+q 没有实根. 置 $x+\frac{p}{2} = u$、$q-\frac{p^2}{4} = a^2$, 得

$$\int \frac{bx+c}{(x^2+px+q)^k}dx = \int \frac{\alpha u+\beta}{(u^2+a^2)^k}du,$$

其中 $\alpha = b, \beta = c - \frac{bp}{2}$.

其次,

$$\begin{aligned} \int \frac{u}{(u^2+a^2)^k}du &= \frac{1}{2}\int \frac{d(u^2+a^2)}{(u^2+a^2)^k} \\ &= \begin{cases} \frac{1}{2(1-k)}(u^2+a^2)^{1-k}, & \text{当} k \neq 1 \text{时}, \\ \frac{1}{2}\ln(u^2+a^2), & \text{当} k = 1 \text{时}. \end{cases} \end{aligned} \tag{10}$$

于是剩下的是处理积分

$$I_k = \int \frac{du}{(u^2+a^2)^k}. \tag{11}$$

分部积分并做初等变换, 有

$$\begin{aligned} I_k &= \int \frac{du}{(u^2+a^2)^k} = \frac{u}{(u^2+a^2)^k} + 2k\int \frac{u^2du}{(u^2+a^2)^{k+1}} \\ &= \frac{u}{(u^2+a^2)^k} + 2k\int \frac{(u^2+a^2)-a^2}{(u^2+a^2)^{k+1}}du = \frac{u}{(u^2+a^2)^k} + 2kI_k - 2ka^2I_{k+1}, \end{aligned}$$

由此得到递推关系式

$$I_{k+1} = \frac{1}{2ka^2}\frac{u}{(u^2+a^2)^k} + \frac{2k-1}{2ka^2}I_k. \tag{12}$$

它使我们能够降低积分 (11) 中的次数 k. 而 I_1 易于算出:

$$I_1 = \int \frac{du}{u^2+a^2} = \frac{1}{a}\int \frac{d\left(\frac{u}{a}\right)}{1+\left(\frac{u}{a}\right)^2} = \frac{1}{a}\operatorname{arctg}\frac{u}{a} + c; \tag{13}$$

从而使用 (12) 和 (13) 就可以算出原函数 (11).

这样, 关于有理函数积分法, 我们证明了下面的

命题 2 *任何有理函数 $R(x) = \dfrac{P(x)}{Q(x)}$ 的原函数都可由有理函数以及超越函数 ln 和 arctg 表出. 原函数的有理部分如果通分, 应该有这样的公分母, 它是多项式 $Q(x)$ 分解出的全部因子的乘积, 只是幂次比在 $Q(x)$ 中少 1.*

例 13 计算 $\displaystyle\int \frac{2x^2+5x+5}{(x^2-1)(x+2)}dx$.

由于被积函数是真分式且它的分母的乘积展开式 $(x-1)(x+1)(x+2)$ 也已知道, 所以我们一下子就求出它的最简分式的和的形式的展开式.

$$\frac{2x^2+5x+5}{(x^2-1)(x+2)} = \frac{A}{x-1} + \frac{B}{x+1} + \frac{C}{x+2}. \tag{14}$$

将 (14) 右端通分, 有

$$\frac{2x^2+5x+5}{(x^2-1)(x+2)} = \frac{(A+B+C)x^2+(3A+B)x+(2A-2B-C)}{(x^2-1)(x+2)}.$$

于是, 让分子的相应的系数相等, 得到方程组

$$\begin{cases} A+B+C=2, \\ 3A+B=5, \\ 2A-2B-C=5, \end{cases}$$

从而求出 $(A,B,C)=(2,-1,1)$.

我们指出, 在所给的情况下, 这些数也可以心算出来. 实际上, 以 $(x-1)$ 遍乘 (14), 然后在所得等式中令 $x=1$, 右边得 A, 而左边是消去分母中的因子 $(x-1)$ 后的分式当 $x=1$ 时的值, 即 $A=\dfrac{2+5+5}{2\times 3}=2$. 类似地可以求出 B 和 c.

于是,

$$\int \frac{2x^2+5x+5}{(x^2-1)(x+2)}dx = 2\int\frac{dx}{x-1}-\int\frac{dx}{x+1}+\int\frac{dx}{x+2}$$
$$=2\ln|x-1|-\ln|x+1|+\ln|x+2|+c$$
$$=\ln\left|\frac{(x-1)^2(x+2)}{x+1}\right|+c.$$

例 14 计算函数 $R(x)$ 的原函数, 这里

$$R(x)=\frac{x^7-2x^6+4x^5-5x^4+4x^3-5x^2-x}{(x-1)^2(x^2+1)^2}.$$

首先指出 $R(x)$ 不是真分式. 因此, 打开括弧求出它的分母 $Q(x)=x^6-2x^5+3x^4-4x^3-2x+1$, 用它来除分子, 得到

$$R(x)=x+\frac{x^5-x^4+x^3-3x^2-2x}{(x-1)^2(x^2+1)^2}.$$

然后再求真分式 $\dfrac{x^5-x^4+x^3-3x^2-2x}{(x-1)^2(x^2+1)^2}$ 的展开式

$$\frac{x^5-x^4+x^3-3x^2-2x}{(x-1)^2(x^2+1)^2}=\frac{A}{(x-1)^2}+\frac{B}{x-1}+\frac{Cx+D}{(x^2+1)^2}+\frac{Ex+F}{x^2+1}. \tag{15}$$

当然, 可以用正规的办法写出六个未知数的六个方程来求出展开式. 但我们不这样做. 我们向大家演示另外一些可能用到的技巧.

我们这样求系数 A, 用 $(x-1)^2$ 遍乘等式 (15), 然后令 $x=1$, 得到 $A=-1$.

代入已知的值 $A=-1$, 把分式 $\dfrac{A}{(x-1)^2}$ 移到等式 (15) 的左边. 那时得到

$$\frac{x^4+x^3+2x^2+x-1}{(x-1)(x^2+1)^2}=\frac{B}{x-1}+\frac{Cx+D}{(x^2+1)^2}+\frac{Ex+F}{x^2+1}, \tag{16}$$

由此, 以 $(x-1)$ 遍乘 (16) 然后令 $x=1$, 求得 $B=1$.

现在把分式 $\dfrac{B}{x-1}$ 移到等式 (16) 的左边, 得

$$\frac{x^2+x+2}{(x^2+1)^2}=\frac{Cx+D}{(x^2+1)^2}+\frac{Ex+F}{x^2+1}. \tag{17}$$

将等式 (17) 右端通分, 让两边分子相等

$$x^2+x+2=Ex^3+Fx^2+(C+E)x+(D+F),$$

由此推出

$$\begin{cases} E = 0, \\ F = 1, \\ C + E = 1, \\ D + F = 2 \end{cases}$$

即 $(C, D, E, F) = (1, 1, 0, 1)$.

现在我们已经知道等式 (15) 中的全部系数. 前两个分式求积时分别给出 $\dfrac{1}{x-1}$ 和 $\ln|x-1|$. 还有

$$\int \frac{Cx + D}{(x^2+1)^2} dx = \int \frac{x+1}{(x^2+1)^2} dx$$
$$= \frac{1}{2} \int \frac{d(x^2+1)}{(x^2+1)^2} + \int \frac{dx}{(x^2+1)^2} = \frac{-1}{2(x^2+1)} + I_2,$$

其中

$$I_2 = \int \frac{dx}{(x^2+1)^2} = \frac{1}{2} \frac{x}{(x^2+1)^2} + \frac{1}{2} \operatorname{arctg} x,$$

这是从 (12) 和 (13) 推出的.

最后

$$\int \frac{Ex + F}{x^2+1} dx = \int \frac{1}{x^2+1} dx = \operatorname{arctg} x.$$

把所有的积分收拢起来, 最后有

$$\int R(x) dx = \frac{1}{2} x^2 + \frac{1}{x-1} + \frac{1}{2} \frac{x}{(x^2+1)^2} + \ln|x-1| + \frac{3}{2} \operatorname{arctg} x + c.$$

现在来考察一些常见的可以归结为求有理函数的原函数的不定积分.

4. $\int R(\cos x, \sin x) dx$ 型的原函数 设 $R(u, v)$ 是 u 和 v 的有理函数, 即多项式 $P(u, v)$ 和 $Q(u, v)$ 的比, 其中 $P(u, v)$ 和 $Q(u, v)$ 是单项式 $u^m v^n (m、n = 0, 1, 2, \cdots)$ 的线性组合.

计算原函数 $\int R(\cos x, \sin x) dx$ 有好几种方法, 其中有一种是完全万能的, 虽然并不总是最省事的.

a. 做替换 $t = \operatorname{tg} \dfrac{x}{2}$. 因为

$$\cos x = \frac{1 - \operatorname{tg}^2 \frac{x}{2}}{1 + \operatorname{tg}^2 \frac{x}{2}}, \quad \sin x = \frac{2 \operatorname{tg} \frac{x}{2}}{1 + \operatorname{tg}^2 \frac{x}{2}},$$
$$dt = \frac{1}{2} \frac{dx}{\cos^2 \frac{x}{2}}, \quad 即\ dx = \frac{2dt}{1 + \operatorname{tg}^2 \frac{x}{2}},$$

所以

$$\int R(\cos x,\sin x)dx=\int R\left(\frac{1-t^2}{1+t^2},\frac{2t}{1+t^2}\right)\frac{2}{1+t^2}dt.$$

于是问题化为对有理函数求积分.

但是这样去做常常导致非常复杂的有理函数, 因此, 应该注意到在许多场合存在着将积分有理化的别的可行的办法.

b. 对于 $\int R(\cos^2 x,\sin^2 x)dx$ 型或 $\int r(\operatorname{tg} x)dx$ 型的积分, 其中 $r(u)$ 是有理函数, 这时用替换 $t=\operatorname{tg} x$ 比较方便, 因为

$$\cos^2 x=\frac{1}{1+\operatorname{tg}^2 x},\sin^2 x=\frac{\operatorname{tg}^2 x}{1+\operatorname{tg}^2 x},$$

$$dt=\frac{dx}{\cos^2 x}\text{即}dx=\frac{dt}{1+\operatorname{tg}^2 x}.$$

经过这一替换, 我们分别得到

$$\int R(\cos^2 x,\sin^2 x)dx=\int R\left(\frac{1}{1+t^2},\frac{t^2}{1+t^2}\right)\frac{dt}{1+t^2},$$

$$\int r(\operatorname{tg} x)dx=\int r(t)\frac{dt}{1+t^2}.$$

c. 对于 $\int R(\cos x,\sin^2 x)\sin xdx$ 或 $\int R(\cos^2 x,\sin x)\cos xdx$ 型的积分, 可以把函数 $\sin x$、$\cos x$ 移到微分号下, 分别做替换 $t=\cos x$ 和 $t=\sin x$. 经这些替换, 积分化为

$$-\int R(t,1-t^2)dt \quad \text{和} \quad \int R(1-t^2,t)dt.$$

例 15

$$\int\frac{dx}{3+\sin x}=\int\frac{1}{3+\dfrac{2t}{1+t^2}}\cdot\frac{2dt}{1+t^2}$$

$$=2\int\frac{dt}{3t^2+2t+3}=\frac{2}{3}\int\frac{d\left(t+\frac{1}{3}\right)}{\left(t+\frac{1}{3}\right)^2+\frac{8}{9}}=\frac{2}{3}\int\frac{du}{u^2+\left(\frac{2\sqrt{2}}{3}\right)^2}$$

$$=\frac{1}{\sqrt{2}}\operatorname{arctg}\frac{3u}{2\sqrt{2}}+c=\frac{1}{\sqrt{2}}\operatorname{arctg}\frac{3t+1}{\sqrt{2}}+c$$

$$=\frac{1}{\sqrt{2}}\operatorname{arctg}\frac{3\operatorname{tg}\frac{x}{2}+1}{2\sqrt{2}}+c.$$

此处我们使用了万能替换 $t=\operatorname{tg}\frac{x}{2}$.

例 16

$$\int\frac{dx}{(\sin x+\cos x)^2}=\int\frac{dx}{\cos^2 x(\operatorname{tg}x+1)^2}$$
$$=\int\frac{d\operatorname{tg}x}{(\operatorname{tg}x+1)^2}=\int\frac{dt}{(t+1)^2}=-\frac{1}{t+1}+c=c-\frac{1}{1+\operatorname{tg}x}.$$

例 17

$$\int\frac{dx}{2\sin^2 3x-3\cos^2 3x+1}$$
$$=\int\frac{dx}{\cos^2 3x[2\operatorname{tg}^2 3x-3+(1+\operatorname{tg}^2 3x)]}=\frac{1}{3}\int\frac{d\operatorname{tg}3x}{3\operatorname{tg}^2 3x-2}$$
$$=\frac{1}{3}\int\frac{dt}{3t^2-2}=\frac{1}{6}\sqrt{\frac{2}{3}}\int\frac{d\sqrt{\frac{3}{2}}t}{\frac{3}{2}t^2-1}=\frac{1}{3\sqrt{6}}\int\frac{du}{u^2-1}$$
$$=\frac{1}{6\sqrt{6}}\ln\left|\frac{u-1}{u+1}\right|+c=\frac{1}{6\sqrt{6}}\ln\left|\frac{\sqrt{\frac{3}{2}}t-1}{\sqrt{\frac{3}{2}}t+1}\right|+c$$
$$=\frac{1}{6\sqrt{6}}\ln\left|\frac{\operatorname{tg}3x-\sqrt{\frac{2}{3}}}{\operatorname{tg}3x+\sqrt{\frac{2}{3}}}\right|+c.$$

例 18

$$\int\frac{\cos^3 x}{\sin^7 x}dx=\int\frac{\cos^2 x d\sin x}{\sin^7 x}$$
$$=\int\frac{1-t^2}{t^7}dt=\int(t^{-7}-t^{-5})dt=-\frac{1}{6}t^{-6}+\frac{1}{4}t^{-4}+c$$
$$=\frac{1}{4\sin^4 x}-\frac{1}{6\sin^6 x}+c.$$

5. $\int R(x,y(x))dx$ 型的原函数 像在第 4 段中那样, 设 $R(x,y)$ 是有理函数. 我们考察形如

$$\int R(x,y(x))dx$$

的某些特殊的原函数, 其中 $y=y(x)$ 是 x 的函数.

首先, 很明白, 如果能够进行替换 $x=x(t)$ 使得两个函数 $x=x(t)$ 和 $y=y(x(t))$ 都是 t 的有理函数的话, 那么 $x'(t)$ 同样也是有理函数, 从而

$$\int R(x,y(x))dx=\int R(x(t),y(x(t)))x'(t)dt,$$

也就是说, 问题归结为有理函数的求积.

我们考虑下列给定函数 $y = y(x)$ 的特殊情形.

a. 如果 $y = \sqrt[n]{\dfrac{ax+b}{cx+d}}$, 其中 $n \in \mathbb{N}$, 那么, 置 $t^n = \dfrac{ax+b}{cx+d}$ 得到 $x = \dfrac{t^n d - b}{a - ct^n}$、$y = t$, 从而被积表达式被有理化.

例 19

$$\begin{aligned}
\int \sqrt[3]{\frac{x-1}{x+1}}dx &= \int t d\left(\frac{t^3+1}{1-t^3}\right) \\
&= t\frac{t^3+1}{1-t^3} - \int \frac{t^3+1}{1-t^3}dt = t\frac{t^3+1}{1-t^3} - \int \left(\frac{2}{1-t^3} - 1\right)dt \\
&= t\frac{t^3+1}{1-t^3} + t - 2\int \frac{dt}{(1-t)(1+t+t^2)} \\
&= t\frac{t^3+1}{1-t^3} + t - 2\int \left(\frac{\frac{1}{3}}{1-t} + \frac{\frac{2}{3}+\frac{1}{3}t}{1+t+t^2}\right)dt \\
&= \frac{2t}{1-t^3} + \frac{2}{3}\ln|1-t| - \frac{2}{3}\int \frac{\left(t+\frac{1}{2}\right)+\frac{3}{2}}{\left(t+\frac{1}{2}\right)^2+\frac{3}{4}}dt \\
&= \frac{2t}{1-t^3} + \frac{2}{3}\ln|1-t| - \frac{1}{3}\ln\left[\left(t+\frac{1}{2}\right)^2+\frac{3}{4}\right] \\
&\quad -\frac{2}{\sqrt{3}}\operatorname{arctg}\frac{2}{\sqrt{3}}\left(t+\frac{1}{2}\right) + c,
\end{aligned}$$

其中 $t = \sqrt[3]{\dfrac{x-1}{x+1}}$.

b. 现在考察 $y = \sqrt{ax^2+bx+c}$ 的情形, 即考虑下型的积分

$$\int R(x, \sqrt{ax^2+bx+c})dx.$$

从三项式中分出完全平方并做适当的线性变量替换, 我们把一般情况化为下列三种简单情形之一:

$$\int R(t, \sqrt{t^2+1})dt, \int R(t, \sqrt{t^2-1})dt, \int R(t, \sqrt{1-t^2})dt. \tag{18}$$

为把这些积分有理化, 只须分别令 (欧拉替换)

$\sqrt{t^2+1} = tu+1$ 或 $\sqrt{t^2+1} = tu-1$ 或 $\sqrt{t^2+1} = t-u$;

$\sqrt{t^2-1} = (t-1)u$ 或 $\sqrt{t^2-1} = (t+1)u$ 或 $\sqrt{t^2-1} = t-u$;

$\sqrt{1-t^2} = u(1-t)$ 或 $\sqrt{1-t^2} = u(1+t)$ 或 $\sqrt{1-t^2} = tu \pm 1$.

这些替换也是由欧拉提出的 (见本节末的问题 3).

作为例子我们来验证经第一个替换之后，把第一个积分化为有理函数的积分. 事实上，若 $\sqrt{t^2+1}=tu+1$，则 $t^2+1=t^2u^2+2tu+1$，由此，

$$t=\frac{2u}{1-u^2}.$$

以及

$$\sqrt{t^2+1}=\frac{1+u^2}{1-u^2}$$

这样一来，t 和 $\sqrt{t^2+1}$ 都通过 u 有理表出了，因而积分化成了有理函数的积分.

积分 (18) 还可经替换 $t=\operatorname{sh}\varphi$、$t=\operatorname{ch}\varphi$、$t=\sin\varphi$ (或 $t=\cos\varphi$) 分别化为三角形式

$$\int R(\operatorname{sh}\varphi,\operatorname{ch}\varphi)\operatorname{ch}\varphi d\varphi,\int R(\operatorname{ch}\varphi,\operatorname{sh}\varphi)\operatorname{sh}\varphi d\varphi$$

以及

$$\int R(\sin\varphi,\cos\varphi)\cos\varphi d\varphi,-\int R(\cos\varphi,\sin\varphi)\sin\varphi d\varphi.$$

例 20

$$\int\frac{dx}{x+\sqrt{x^2+2x+2}}=\int\frac{dx}{x+\sqrt{(x+1)^2+1}}=\int\frac{dt}{t-1+\sqrt{t^2+1}}.$$

令 $\sqrt{t^2+1}=u-t$，有$1=u^2-2tu$，由此 $t=\dfrac{u^2-1}{2u}$. 因此

$$\begin{aligned}\int\frac{dt}{t-1+\sqrt{t^2+1}}&=\frac{1}{2}\int\frac{1}{u-1}\left(1+\frac{1}{u^2}\right)du=\frac{1}{2}\int\frac{du}{u-1}+\frac{1}{2}\int\frac{du}{u^2(u-1)}\\&=\frac{1}{2}\ln|u-1|+\frac{1}{2}\int\left(\frac{1}{u-1}-\frac{1}{u^2}-\frac{1}{u}\right)dw\\&=\frac{1}{2}\ln|u-1|+\frac{1}{2}\ln\left|\frac{u-1}{u}\right|+\frac{1}{2u}+c.\end{aligned}$$

现在只要再做相反的替换 $u=t+\sqrt{t^2+1}$ 和 $t=x+1$ 即可.

c. 椭圆积分. 形如

$$\int R(x,\sqrt{P(x)})dx \tag{19}$$

的原函数也是很重要的，其中 $P(x)$ 是次数 $n>2$ 的多项式. 阿贝尔和刘维尔曾经证明，这样的积分一般说来已不能用初等函数表示.

当 $n=3$ 和 $n=4$ 时，积分 (19) 叫做椭圆积分，而当 $n>4$ 时叫做超椭圆积分.

可以证明, 一般的椭圆积分经过初等替换, 在不计一个初等函数被加项的条件下, 可归结为下面三个标准椭圆积分:

$$\int \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}, \tag{20}$$

$$\int \frac{x^2dx}{\sqrt{(1-x^2)(1-k^2x^2)}}, \tag{21}$$

$$\int \frac{dx}{(1+hx^2)\sqrt{(1-x^2)(1-k^2x^2)}}, \tag{22}$$

其中 h 和 k 是参数, 且在三种情况下参数 k 都位于开区间 $]0,1[$ 之中.

经替换 $x=\cos\varphi$, 这些积分可化为下列典则积分或它们的线性组合:

$$\int \frac{d\varphi}{\sqrt{1-k^2\sin\varphi^2}}, \tag{23}$$

$$\int \sqrt{1-k^2\sin^2\varphi}d\varphi, \tag{24}$$

$$\int \frac{d\varphi}{(1-h\sin^2\varphi)\sqrt{1-k^2\sin^2\varphi}}. \tag{25}$$

积分 (23)、(24)、(25) 分别叫作 (拉格朗日形式的) *第一类、第二类和第三类椭圆积分*. 用 $F(k,\varphi)$ 和 $E(k,\varphi)$ 分别代表由条件 $F(k,0)=0$ 和 $E(k,0)=0$ 确定的第一类椭圆积分和第二类椭圆积分 (23) 和 (24).

函数 $F(k,\varphi)$ 和 $E(k,\varphi)$ 常常被用到, 因此对于 $0<k<1$ 和 $0\leqslant\varphi\leqslant\pi/2$ 为它们编制了相当详细的函数值表.

阿贝尔指出, 在复数域中很自然地看到椭圆积分与所谓椭圆函数的不可分割的联系. 椭圆函数与椭圆积分的关系, 举例来说, 就如同函数 $\sin\varphi$ 与积分 $\int \frac{d\varphi}{\sqrt{1-\varphi^2}} = \arcsin\varphi$ 的关系一样.

练 习

1. *奥斯特洛格拉德斯基*[①]*方法*, 这是一个将有理真分式的不定积分的有理部分划分出来的方法. 设 $\frac{P(x)}{Q(x)}$ 是有理真分式; $q(x)$ 是与 $Q(x)$ 有相同的根但都是一重根的多项式; $Q_1(x)=\frac{Q(x)}{q(x)}$.

   证明

   a) 下面的奥斯特洛格拉德斯基公式成立:

$$\int \frac{P(x)}{Q(x)}dx = \frac{P_1(x)}{Q_1(x)} + \int \frac{p(x)}{q(x)}dx, \tag{26}$$

①奥斯特洛格拉德斯基 (М. В. Остроградский)(1801—1861) —— 杰出的俄罗斯力学家和数学家, 彼得堡数学学派中实际应用研究方向的倡导者之一.

其中 $\dfrac{P_1(x)}{Q_1(x)}$、$\dfrac{p(x)}{q(x)}$ 是有理真分式, 且 $\displaystyle\int \frac{p(x)}{q(x)}dx$ 是超越函数.

(根据这个结果,(26) 中的分式 $\dfrac{P_1(x)}{Q_1(x)}$叫作积分$\displaystyle\int \frac{P(x)}{Q(x)}dx$ 的有理部分.)

b) 在微分奥斯特洛格拉德斯基公式所得到的公式

$$\frac{P(x)}{Q(x)} = \left[\frac{P_1(x)}{Q_1(x)}\right]' + \frac{p(x)}{q(x)}$$

中, 分式 $\left[\dfrac{P_1(x)}{Q_1(x)}\right]'$ 经适当约简可得分母 $Q(x)$.

c) 既使不知道多项式 $Q(x)$ 的根, 也可以用代数的办法求出多项式 $q(x)$、$Q_1(x)$ 并随后求出多项式 $p(x)$、$P_1(x)$. 这样一来, 既使不计算整个原函数, 也完全可以求出积分 (26) 的有理部分.

d) 如果 $P(x) = 2x^6 + 3x^5 + 6x^4 + 6x^3 + 10x^2 + 3x + 2$, 而

$$Q(x) = x^7 + 3x^6 + 5x^5 + 7x^4 + 7x^3 + 5x^2 + 3x + 1,$$

试分出积分 (26) 的有理部分 (见本章 §5 例 17).

2. 设欲求原函数

$$\int R(\cos x, \sin x)dx, \tag{27}$$

其中 $R(u,v) = \dfrac{P(u,v)}{Q(u,v)}$ 是有理函数.

证明:

a) 若 $R(-u,v) = R(u,v)$, 则 $R(u,v)$ 形如 $R_1(u^2,v)$.

b) 若 $R(-u,v) = -R(u,v)$, 则 $R(u,v) = R_2(u^2,v)u$, 且替换 $t = \sin x$ 把积分 (27) 有理化.

c) 若 $R(-u,-v) = R(u,v)$, 则 $R(u,v) = R_3\left(\dfrac{u}{v}, v^2\right)$, 且替换 $t = \operatorname{tg} x$ 把积分 (27) 有理化.

3. $\displaystyle\int R\left(x, \sqrt{ax^2+bx+c}\right)dx$型积分

a) 验证: 经下列欧拉替换

$$t = \sqrt{ax^2+bx+c} \pm \sqrt{a}x, \text{若} a > 0,$$
$$t = \sqrt{\frac{x-x_1}{x-x_2}}, \text{若} x_1 \text{、} x_2 \text{是三项式} ax^2+bx+c \text{的实根},$$

积分

$$\int R\left(x, \sqrt{ax^2+bx+c}\right)dx \tag{28}$$

被化为有理函数的积分.

b) 设 (x_0, y_0) 是曲线 $y^2 = ax^2 + bx + c$ 上的点, 而 t 是过点 (x_0, y_0) 与曲线在点 (x, y) 处相交的直线的斜率. 通过 (x_0, y_0) 和 t 来表示坐标 (x, y) 并把这些公式与欧拉替换联系起来.

c) 由代数方程 $P(x,y)=0$ 确定的曲线, 如果能借助于有理函数 $x(t)$、$y(t)$ 用参数形式 $x=x(t)$、$y=y(t)$ 表示的话, 就叫做有理曲线. 证明: 若 $R(u,v)$ 是有理函数, $y(x)$ 是满足给出有理曲线的方程 $P(x,y)=0$ 的一个代数函数, 那么 $\int R(x,y(x))dx$ 归结为有理函数的积分.

d) 证明: 永远可以把积分 (28) 归结为计算下列三种类型的积分:

$$\int \frac{P(x)}{\sqrt{ax^2+bx+c}}dx, \int \frac{dx}{(x-x_0)^k\sqrt{ax^2+bx+c}},$$
$$\int \frac{(Ax+B)dx}{(x^2+px+q)^m\sqrt{ax^2+bx+c}}.$$

4. a) 试证: 微分二项式的积分 $\int x^m(a+bx^n)^p dx$(其中 m,n,p 是有理数) 可归结为积分

$$\int (a+bt)^p t^q dt, \tag{29}$$

其中 p、q 是有理数.

b) 如果 p, q、$p+q$ 这三个数中有一个是整数的话, 积分 (29) 就可以用初等函数表示.(切比雪夫证明, 除了这种情形, 积分 (29) 不能用初等函数表示.)

5. 椭圆积分

a) 任何实系数三次多项式都有实根 x_0, 且替换 $x-x_0=t^2$ 导致形如 $t^2(at^4+bt^3+ct^2+dt+e)$ 的多项式, 其中 $a\neq 0$.

b) 函数 $R(x,\sqrt{P(x)})$(其中 $R(u,v)$ 是有理函数, 而 P 是三次或四次多项式), 可以化成 $R_1(t,\sqrt{at^4+bt^3+ct^2+dt+e})$, 其中 $a\neq 0$.

c) 四次多项式 $ax^4+bx^3+\cdots+e$ 可表成乘积

$$a(x^2+p_1x+q_1)(x^2+p_2x+q_2),$$

且经替换 $x=\dfrac{\alpha t+\beta}{\gamma t+1}$ 总可化成

$$\frac{(M_1+N_1t^2)(M_2+N_2t^2)}{(\gamma t+1)^2}.$$

d) 函数 $R(x,\sqrt{ax^4+bx^3+\cdots+e})$ 经替换 $x=\dfrac{\alpha t+\beta}{\gamma t+1}$ 可化成

$$R_1(t,\sqrt{A(1+m_1t^2)(1+m_2t^2)}).$$

e) 函数 $R(x,\sqrt{y})$ 可以表成和 $R_1(x,y)+\dfrac{R_2(x,y)}{\sqrt{y}}$, 其中 R_1 和 R_2 是有理函数.

f) 任何有理函数都可以表成偶的和奇的两个有理函数的和.

g) 若有理函数 $R(x)$ 是偶的, 则有 $r(x^2)$ 的形式, 而若它是奇的, 则有 $xr(x^2)$ 的形式, 其中 $r(x)$ 是有理函数.

h) 任何函数 $R(x,\sqrt{y})$ 都可化为

$$R_1(x,y)+\frac{R_2(x^2,y)}{\sqrt{y}}+\frac{R_3(x^2,y)}{\sqrt{y}}x.$$

i) 任何 $\int R(x,\sqrt{P(x)})dx$ 型的积分 (其中 $P(x)$ 是四次多项式) 都可以化成积分

$$\int\frac{r(t^2)dt}{\sqrt{A(1+m_1t^2)(1+m_2t^2)}}$$

再加上一个初等的项, 其中 $r(t)$ 是有理函数, $A=\pm1$.

j) 如果 $|m_1|>|m_2|>0$, 那么

$$\sqrt{|m_1|}t=x, \sqrt{|m_1|}t=\sqrt{1-x^2}, \sqrt{|m_1|}t=\frac{x}{\sqrt{1-x^2}},$$
$$\sqrt{|m_1|}t=\frac{1}{\sqrt{1-x^2}};$$

这四种形式的变换中有一种可把积分 $\int\frac{r(t^2)dt}{\sqrt{A(1+m_1t^2)(1+m_2t^2)}}$ 化为

$$\int\frac{\tilde{r}(x^2)dx}{\sqrt{(1-x^2)(1-k^2x^2)}},$$

其中 $0<k<1$ 而 $\tilde{r}$ 是有理函数.

k) 求出使积分

$$\int\frac{x^{2n}dx}{\sqrt{(1-x^2)(1-k^2x^2)}}, \int\frac{dx}{(x^2-a)^m\sqrt{(1-x^2)(1-k^2x^2)}}$$

中指数 $2n$、m 降低的公式.

l) 任何椭圆积分

$$\int R(x,\sqrt{P(x)})dx$$

(其中 P 是四次多项式) 都可以化为 (20)、(21)、(22) 这三种标准的积分中的一种加上一个初等函数.

m) 把积分 $\int\frac{dx}{\sqrt{1+x^3}}$ 通过标准椭圆积分表示出来.

n) 用椭圆积分表出函数 $\frac{1}{\sqrt{\cos 2x}}$ 和 $\frac{1}{\sqrt{\cos\alpha-\cos x}}$ 的原函数.

6. 求下列非初等的特殊函数的原函数, 精确到线性函数 $Ax+B$:

a) $\mathrm{Ei}(x)=\int\frac{e^x}{x}dx$(积分指数).

b) $\mathrm{Si}(x)=\int\frac{\sin x}{x}dx$(积分正弦).

c) $\mathrm{Ci}(x)=\int\frac{\cos x}{x}dx$(积分余弦).

d) $\mathrm{Shi}(x)=\int\frac{\operatorname{sh} x}{x}dx$(积分双曲正弦).

e) $\mathrm{Chi}(x) = \int \frac{\mathrm{ch}\, x}{x} dx$(积分双曲余弦).

f) $S(x) = \int \sin x^2 dx$(菲涅耳积分).

g) $C(x) = \int \cos x^2 dx$(菲涅耳积分).

h) $\Phi(x) = \int e^{-x^2} dx$(欧拉-泊松积分).

i) $\mathrm{li}(x) = \int \frac{dx}{\ln x}$(积分对数).

7. 验证, 精确到常数下列等式成立:

a) $\mathrm{Ei}(x) = \mathrm{li}(e^x)$,

b) $\mathrm{Chi}(x) = \frac{1}{2}[\mathrm{Ei}(x) + \mathrm{Ei}(-x)]$,

c) $\mathrm{Shi}(x) = \frac{1}{2}[\mathrm{Ei}(x) - \mathrm{Ei}(-x)]$,

d) $\mathrm{Ei}(ix) = \mathrm{Ci}(x) + i\, \mathrm{Si}(x)$,

e) $e^{i\pi/4}\Phi(xe^{-i\pi/4}) = C(x) + iS(x)$.

8. 形如

$$\frac{dy}{dx} = \frac{f(x)}{g(y)}$$

的微分方程叫作*变量分离的方程*, 因为可以把它改写成

$$g(y)dy = f(x)dx,$$

这里变量 x 和 y 已被分开. 分开变量之后就可以分别计算原函数而解出方程:

$$\int g(y)dy = \int f(x)dx + c.$$

试解方程:

a) $2x^3yy' + y^2 = 2$.

b) $xyy' = \sqrt{1+x^2}$.

c) $y' = \cos(y+x)$, 令 $u(x) = y(x) + x$.

d) $x^2y' - \cos 2y = 1$, 并挑出当 $x \to +0$ 时满足条件 $y(x) \to 0$ 的解.

e) $\frac{1}{x}y'(x) = \mathrm{Si}(x)$.

f) $\frac{y'(x)}{\cos x} = C(x)$.

9. 跳伞者从 1.5km 高处跳下, 而降落伞在 0.5km 高处张开. 试问, 从跳下至开伞经历了多少时间? 在正常密度空气中人落下的极限速度取作 50m/s. 假定空气阻力正比于:

a) 速度;

b) 速度的平方,

(压强随高度的变化忽略不计) 分别解答这个问题.

10. 已经知道水从容器底部的小孔流出的速度可以相当精确地按公式 $0.6\sqrt{2gH}$ 来计算, 其中 g 是重力加速度而 H 是孔以上水面的高度.

底部有小孔的圆柱形桶垂直地放着. 5min 从整桶水中流走一半的水, 问需多少时间流走全部的水?

11. 当水从旋转体容器中流出时, 为了使得水面均匀下降, 该容器应具有怎样的形状?(参看问题 10)

12. 在容积为 10^4m^3 的厂房中, 通风机在 1min 内送入 10^3m^3 含 0.04%CO_2 的新鲜空气, 同时从厂房中抽走同样数量的混合气体. 早上九点钟职工上班, 经过半个钟头空气中的 CO_2 增加到 0.12%. 问经过两小时后厂房中含有多少二氧化碳气体?

第六章　积分

§1. 积分定义和可积函数集的描述

1. 问题和启发性想法　设有一点沿数轴运动, 它在时刻 t 的坐标是 $s(t)$, 而 $v(t) = s'(t)$ 是它在同一时刻 t 的速度. 假定我们知道这个点在时刻 t_0 的位置是 $s(t_0)$, 而且还给我们提供了它的速度的数据. 我们想用这些资料, 对任意确定的时刻 $t > t_0$ 计算 $s(t)$.

如果认为速度 $v(t)$ 是连续变化的, 那么点在一小段时间内的位移将近似地等于该时间间隔内任一时刻 τ 的速度 $v(\tau)$ 与这一小段时间间隔的值 Δt 的乘积 $v(\tau)\Delta t$. 这提示我们, 指定一些时刻 $t_i(i = 0, \cdots, n), t_0 < t_1 < \cdots < t_n = t$, 以分割区间 $[t_0, t]$. 设间隔 $[t_{i-1}, t_i]$ 都取得很小. 记 $\Delta t_i = t_i - t_{i-1}$, 取 $\tau_i \in [t_{i-1}, t_i]$. 则有近似等式

$$s(t) - s(t_0) \approx \sum_{i=1}^{n} v(\tau_i)\Delta t_i.$$

可以想象, 如果对区间 $[t_0, t]$ 所做的分割越细, 这个近似等式就越准确. 这样一来, 可以设想, 在当分划的最大间隔长 λ 趋于零的极限情形下, 我们将得到精确等式

$$\lim_{\lambda \to 0} \sum_{i=1}^{n} v(\tau_i)\Delta t_i = s(t) - s(t_0). \tag{1}$$

这个等式不是别的, 正是全部分析学的基本公式 — 牛顿-莱布尼茨公式. 一方面, 它使我们能够根据导数 $v(t)$ 求出它的原函数数值; 另一方面, 又可根据已经用某种方法得到的函数 $v(t)$ 的原函数 $s(t)$, 求出位于等式左边的和式 $\sum\limits_{i=1}^{n} v(\tau_i)\Delta t_i$ 的极限.

在许多极其不同的场合都会遇到这种和式, 我们称它为*积分和*.

譬如, 我们试图仿效阿基米德方法求位于抛物线 $y = x^2$ 下边、闭区间 $[0,1]$ 上边的图形的面积 (图 47). 这里不详细讨论图形的面积概念, 以后会论述它. 像阿基米德一样, 我们将利用已经会计算其面积的最简单的图形 —— 矩形, 对上述图形施行穷尽法. 取点 $0 = x_0 < x_1 < \cdots < x_n = 1$, 把区间 $[0,1]$ 分成一些小区间 $[x_{i-1}, x_i]$. 显然, 我们可以把图上画出的那些矩形的面积的和作为要求的面积 σ 的近似值:

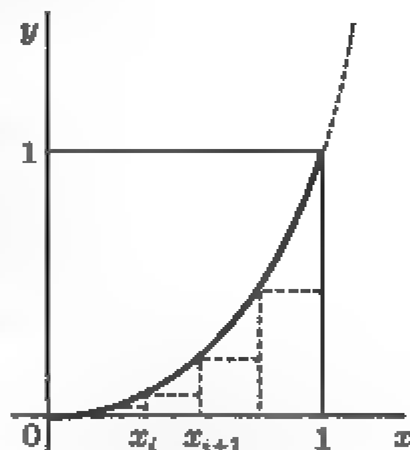


图 47

$$\sigma \approx \sum_{i=1}^{n} x_{i-1}^2 \Delta x_i;$$

这里 $\Delta x_i = x_i - x_{i-1}$. 设 $f(x) = x^2$ 以及 $\xi_i = x_{i-1}$. 我们把所得的公式改写成如下形式:

$$\sigma \approx \sum_{i=1}^{n} f(\xi_i)\Delta x_i.$$

使用这些记号, 当过渡到极限时就得到

$$\lim_{\lambda \to 0} \sum_{i=1}^{n} f(\xi_i)\Delta x_i = \sigma. \tag{2}$$

同上边一样, 这里的 λ 是分划的最大区间长.

公式 (2) 只在记号上与公式 (1) 不同. 暂时不管 $f(\xi_i)$ 和 Δx_i 的几何意义, 而把 x 看成时间, $f(x)$ 看成速度, 来求 $f(x)$ 的原函数 $F(x)$. 然后, 根据公式 (1) 将得到 $\sigma = F(1) - F(0)$.

在我们的情形, $f(x) = x^2$, 因此

$$F(x) = \frac{1}{3}x^3 + c, \sigma = F(1) - F(0) = 1/3.$$

这正是阿基米德的结果, 而他是用直接计算 (2) 中的极限得到的.

积分和的极限叫做*积分*. 这样, 牛顿–莱布尼茨公式(1) 就把积分与原函数联系起来了.

现在我们来精确地叙述并检验上边那些从一般想法得到的具有启发性的东西.

2. 黎曼积分的定义

a. 分划

定义 1 闭区间 $[a,b](a < b)$ 的*分划* P 指的是由这个区间的有限多个点 $x_0, \cdots, x_n$ 做成的点组, 其中 $a = x_0 < x_1 < \cdots < x_n = b$.

区间 $[x_{i-1}, x_i](i = 1, \cdots, n)$ 叫做*分划 P 的区间*.

分划 P 的最大区间长 $\lambda(P)$ 叫做*分划 P 的参数*.

定义 2 如果 P 是闭区间 $[a,b]$ 上的一个分划, 而且在它的每个区间 $[x_{i-1},x_i]$ 中选定了一个点 $\xi_i\in[x_{i-1},x_i](i=1,\cdots,n)$, 则说给出了区间 $[a,b]$ 的一个*带标志点的分划* (P,ξ).

数组 $(\xi_1,\cdots,\xi_n)$ 用一个符号 ξ 表示.

b. 分划集的基 给定闭区间 $[a,b]$, 在它的带标志点的分划的集合 $\mathcal{P}$ 中考察如下的基 $\mathcal{B}=\{B_d\}$. 基 $\mathcal{B}$ 的元素 $B_d(d>0)$ 是区间 $[a,b]$ 的一切满足条件 $\lambda(P)<d$ 的带标志点的分划 (P,ξ) 的集合.

我们来验证 $\{B_d\}(d>0)$ 确实是 $\mathcal{P}$ 的基.

首先, $B_d\neq\varnothing$. 事实上, 对任何数 $d>0$, 显然存在区间 $[a,b]$ 的分划 P, 其参数 $\lambda(P)<d$(例如, 分成 n 个全等区间的分划). 由此, 也存在带标志点的分划 (P,ξ) 且 $\lambda(P)<d$.

其次, 如果 $0<d_1,0<d_2$ 以及 $d=\min\{d_1,d_2\}$, 则显然有 $B_{d_1}\cap B_{d_2}=B_d\in\mathcal{B}$.

因此, $\mathcal{B}=\{B_d\}$ 确实是 $\mathcal{P}$ 的基.

c. 积分和

定义 3 如果函数 f 在闭区间 $[a,b]$ 上定义, 而 (P,ξ) 是这个区间的一个带标志点的分划, 则和

$$\sigma(f;P,\xi):=\sum_{i=1}^{n}f(\xi_i)\Delta x_i \tag{3}$$

叫做函数 f 在区间 $[a,b]$ 上对应于带标志点的分划 (P,ξ) 的*积分和*, 其中 $\Delta x_i=x_i-x_{i-1}$.

这样, 对于确定的函数 f, 积分和 $\sigma(f;P,\xi)$ 是定义在集合 $\mathcal{P}$ 上的函数 $\Phi(p)=\sigma(f;p)$, 这里 $p=(P,\xi)$ 跑遍区间 $[a,b]$ 上带标志点的分划的集合 $\mathcal{P}$.

由于在 $\mathcal{P}$ 中有基 $\mathcal{B}$, 所以可以提出函数 $\Phi(p)$ 关于这个基的极限的问题.

d. 黎曼积分 设 f 是给定在闭区间 $[a,b]$ 上的函数.

定义 4 称数 I 是函数 f 在闭区间 $[a,b]$ 上的*黎曼积分*, 如果对于任何 $\varepsilon>0$ 可以找到 $\delta>0$, 使对区间 $[a,b]$ 的任何带标志点的分划 (P,ξ), 只要其参数 $\lambda(P)<\delta$, 就有

$$\left|I-\sum_{i=1}^{n}f(\xi_i)\Delta x_i\right|<\varepsilon.$$

因为满足 $\lambda(P)<\delta$ 的那些分划 $p=(P,\xi)$ 组成了上面在带标志点的分划的集合 $\mathcal{P}$ 中引进的基 $\mathcal{B}$ 的元素 B_δ, 所以定义 4 等价于

$$I=\lim_{\mathcal{B}}\Phi(p), \tag{4}$$

亦即: 积分I 是函数 f 在区间 $[a,b]$ 上, 对应于带标志点的分划的积分和的值关于基 $\mathcal{B}$ 的极限.

用符号 $\lambda(P)\to 0$ 表示基 $\mathcal{B}$ 是自然的, 从而, 积分定义可改写成

$$I=\lim_{\lambda(P)\to 0}\sum_{i=1}^{n} f(\xi_i)\Delta x_i$$

的形式.

函数 $f(x)$ 在区间 $[a,b]$ 上的积分用符号

$$\int_a^b f(x)dx$$

表示, 数 a,b 分别叫做积分的下限和上限; f 叫做被积函数, $f(x)dx$ 叫做被积表达式, x 叫做积分变量.

于是,

$$\boxed{\int_a^b f(x)dx := \lim_{\lambda(P)\to 0}\sum_{i=1}^{n} f(\xi_i)\Delta x_i.} \tag{5}$$

定义 5 称函数 f 是闭区间 $[a,b]$ 上的黎曼可积函数, 如果对于它, 公式 (5) 中的积分和当 $\lambda(P)\to 0$ 时的极限存在 (也就是说, 它的黎曼积分有定义).

在区间 $[a,b]$ 上的一切黎曼可积函数所成的集合以 $\mathcal{R}[a,b]$ 表示.

由于暂时不研究黎曼积分以外的其他积分, 为简单起见, 我们约定, 分别以《积分》和《可积函数》代替《黎曼积分》和《黎曼可积函数》这两个术语.

3. 可积函数集 根据积分的定义 (定义 4) 和它另外的形如 (4)、(5) 的表述, 积分是一个特定函数 $\Phi(p)=\sigma(f;P,\xi)$ 的极限, 这个函数是定义在区间 $[a,b]$ 的带标志点的分划 $p=(P,\xi)$ 的集 $\mathcal{P}$ 上的积分和. 这个极限是关于 $\mathcal{P}$ 的基 $\mathcal{B}$ 取的, 我们以 $\lambda(P)\to 0$ 表示它.

这样一来, 函数 f 在闭区间 $[a,b]$ 上的可积性取决于上述极限的存在.

根据柯西准则, 这个极限存在, 当且仅当对于任何 $\varepsilon>0$ 可以找到元素 $B_\delta\in\mathcal{B}$, 使在其中任何点 p',p'' 都成立关系

$$|\Phi(p')-\Phi(p'')|<\varepsilon.$$

更详细地说, 这意味着: 对于任何 $\varepsilon>0$, 可以找到 $\delta>0$, 使对区间 $[a,b]$ 的任何带标志点的分划 $(P',\xi'),(P'',\xi'')$, 只要 $\lambda(P')<\delta$ 和 $\lambda(P'')<\delta$, 就成立不等式

$$|\sigma(f;P',\xi')-\sigma(f;P'',\xi'')|<\varepsilon$$

或

$$\left|\sum_{i=1}^{n'} f(\xi_i')\Delta x_i' - \sum_{i=1}^{n''} f(\xi_i'')\Delta x_i''\right| < \varepsilon. \tag{6}$$

利用柯西准则, 我们将首先得到函数黎曼可积的一个简单的必要条件, 然后得到一些充分条件.

a. 可积性的必要条件

命题 1 *为使定义在闭区间 $[a,b]$ 上的函数 f 在这个区间上黎曼可积, 它必须在这个区间上有界.*

简单地说, 即

$$(f \in \mathcal{R}[a,b]) \Rightarrow (f在[a,b]上有界).$$

◀ 如果 f 在 $[a,b]$ 上无界, 则对区间 $[a,b]$ 的任一分划 P, f 至少在它的一个区间 $[x_{i-1},x_i]$ 上无界. 因此, 可以选出点 $\xi_i \in [x_{i-1},x_i]$ 使 $|f(\xi_i)\Delta x_i|$ 任意大. 于是, 只需改变点 ξ_i 在这个区间中的位置, 就能使积分和 $\sigma(f;P,\xi) = \sum\limits_{i=1}^{n} f(\xi_i)\Delta x_i$ 的绝对值任意大.

在这种情形下, 积分和显然没有有限极限, 同时, 由柯西准则看出, 这时关系式 (6) 甚至对无论多么细的分划都不全成立. ▶

将会看到, 这个必要条件与可积性的必要充分条件还相距很远. 但是, 它告诉我们今后只须研究有界函数.

b. 可积性的充分条件和最重要的可积函数类 首先介绍一些下面用到的记号和知识.

我们约定, 当给定了闭区间 $[a,b]$ 的一个分划 P

$$a = x_0 < x_1 < x_2 < \cdots < x_n = b$$

时, 与表示差 $x_i - x_{i-1}$ 的符号 Δx_i 一起, 我们还使用记号 Δ_i 表示区间 $[x_{i-1},x_i]$.

如果区间 $[a,b]$ 的分划 $\tilde{P}$ 是由分划 P 添加一些新的分点得到的, 则说分划 $\tilde{P}$ 是分划 P 的*开拓*.

当把分划 P 开拓成 $\tilde{P}$ 时, 分划 P 的某些 (可以是其全部) 区间 $\Delta_i = [x_{i-1},x_i]$ 被分割:$x_{i-1} = x_{i0} < \cdots < x_{in_i} = x_i$. 由于这个原因. 对 $\tilde{P}$ 的点用两个指标编号比较方便. 在 x_{ij} 中的第一个指标表示 $x_{ij} \in \Delta_i$, 而第二个指标是在区间 Δ_i 上点的顺序号码. 现在, 自然地应记

$$\Delta x_{ij} := x_{ij} - x_{ij-1} \quad 和 \quad \Delta_{ij} := [x_{ij-1}, x_{ij}].$$

这样一来, $\Delta x_i = \Delta x_{i1} + \cdots + \Delta x_{in_i}$.

由分划 P' 和分划 P'' 的点的并得到的分划 $\tilde{P} = P' \cup P''$, 是同时是分划 P' 和 P'' 的开拓的例子.

最后, 我们提醒注意. 如同过去一样, 符号 $\omega(f, E)$ 表示函数 f 在集合 E 上的振幅, 即

$$\omega(f; E) := \sup_{x', x'' \in E} |f(x') - f(x'')|.$$

特别地, $\omega(f; \Delta_i)$ 是函数 f 在区间 Δ_i 上的振幅. 如果 f 是有界函数, 这个振幅显然是有限的.

现在, 我们叙述并证明

命题 2 *为使闭区间 $[a, b]$ 上的有界函数 f 在这个区间上可积, 只要对于任意的 $\varepsilon > 0$, 都存在 $\delta > 0$, 使对区间 $[a, b]$ 的参数 $\lambda(P) < \delta$ 的分划 P 总成立关系*

$$\left|\sum_{i=1}^{n} \omega(f; \Delta_i) \Delta x_i\right| < \varepsilon.$$

◀ 设 P 是区间 $[a, b]$ 的一个分划, $\tilde{P}$ 是分划 P 的开拓. 我们来估计积分和之差 $\sigma(f; \tilde{P}, \tilde{\xi}) - \sigma(f; P, \xi)$. 利用上面引进的符号, 有

$$\begin{aligned}
|\sigma(f; \tilde{P}, \tilde{\xi}) - \sigma(f; P, \xi)| &= \left|\sum_{i=1}^{n}\sum_{j=1}^{n_i} f(\xi_{ij}) \Delta x_{ij} - \sum_{i=1}^{n} f(\xi_i) \Delta x_i\right| \\
&= \left|\sum_{i=1}^{n}\sum_{j=1}^{n_i} f(\xi_{ij}) \Delta x_{ij} - \sum_{i=1}^{n}\sum_{j=1}^{n_i} f(\xi_i) \Delta x_{ij}\right| \\
&= \left|\sum_{i=1}^{n}\sum_{j=1}^{n_i} (f(\xi_{ij}) - f(\xi_i)) \Delta x_{ij}\right| \\
&\leqslant \sum_{i=1}^{n}\sum_{j=1}^{n_i} |f(\xi_{ij}) - f(\xi_i)| \Delta x_{ij} \\
&\leqslant \sum_{i=1}^{n}\sum_{j=1}^{n_i} \omega(f; \Delta_i) \Delta x_{ij} = \sum_{i=1}^{n} \omega(f; \Delta_i) \Delta x_i.
\end{aligned}$$

在这些计算中, 我们利用了

$$\Delta x_i = \sum_{j=1}^{n_i} \Delta x_{ij} \text{ 以及 } |f(\xi_{ij}) - f(\xi_i)| \leqslant \omega(f; \Delta_i),$$

因为 $\xi_{ij} \in \Delta_{ij} \subset \Delta_i, \xi_i \in \Delta_i$.

从得到的积分和之差的估计推出: 如果函数 f 满足在命题 2 中说的充分条件, 则对每个 $\varepsilon > 0$ 都能找到 $\delta > 0$ 使得对于区间 $[a, b]$ 的参数 $\lambda(P) < \delta$ 的分划 P 以及

P 的开拓 $\tilde{P}$, 关于任意取的标志点 ξ 和 $\tilde{\xi}$, 都有

$$|\sigma(f;\tilde{P},\tilde{\xi})-\sigma(f;P,\xi)|<\frac{\varepsilon}{2}.$$

如果 (P',ξ') 和 (P'',ξ'') 是区间 $[a,b]$ 的任意两个带标志点的分划, 且 $\lambda(P')<\delta, \lambda(P'')<\delta$, 考察分划 $\tilde{P}=P'\cup P''$, 它是两个分划 P', P'' 的开拓, 那么, 根据已经证明的事实, 有

$$|\sigma(f;\tilde{P},\tilde{\xi})-\sigma(f;P',\xi')|<\frac{\varepsilon}{2},$$
$$|\sigma(f;\tilde{P},\tilde{\xi})-\sigma(f;P'',\xi'')|<\frac{\varepsilon}{2}.$$

由此得出, 一旦 $\lambda(P')<\delta, \lambda(P'')<\delta$, 就有

$$|\sigma(f;P',\xi')-\sigma(f;P'',\xi'')|<\varepsilon.$$

这样一来, 根据柯西准则, 积分和的极限

$$\lim_{\lambda(P)\to 0}\sum_{i=1}^{n} f(\xi_i)\Delta x_i$$

存在, 亦即 $f\in\mathcal{R}[a,b]$. ►

推论 1 $(f\in C[a,b])\Rightarrow(f\in\mathcal{R}[a,b])$, *亦即闭区间上的任何连续函数在这个区间上可积.*

◄ 如果函数在闭区间上连续, 则它在这个区间上有界, 从而可积性的必要条件成立. 但是, 闭区间上的连续函数也一致连续, 因此, 对于任何 $\varepsilon>0$ 可以找到 $\delta>0$, 使在任何区间 $\Delta\subset[a,b]$ 上, 只要 Δ 的长小于 δ, 就有 $\omega(f;\Delta)<\dfrac{\varepsilon}{b-a}$. 于是, 对于任何分划 P, 只要其参数 $\lambda(P)<\delta$, 就有

$$\sum_{i=1}^{n}\omega(f;\Delta)\Delta x_i<\frac{\varepsilon}{b-a}\sum_{i=1}^{n}\Delta x_i=\frac{\varepsilon}{b-a}(b-a)=\varepsilon.$$

根据命题 2 可以得到 $f\in\mathcal{R}[a,b]$. ►

推论 2 *如果定义在闭区间 $[a,b]$ 上的有界函数 f 在该区间上最多除有限多个点外是连续的, 则 $f\in\mathcal{R}[a,b]$.*

◄ 设 $\omega(f;[a,b])\leqslant C<\infty$, 而 f 在 $[a,b]$ 上有 k 个间断点. 我们来验证函数 f 满足可积性的充分条件.

对于给定的 $\varepsilon>0$, 取数 $\delta_1=\dfrac{\varepsilon}{8C\cdot k}$ 并做函数 f 的每个间断点的 δ_1–邻域. 这些邻域的并关于区间 $[a,b]$ 的补集是由有限个区间组成的, 在每个这样的区间上 f 都是

连续的, 从而是一致连续的. 由于这种区间只有有限多个, 对 $\varepsilon > 0$ 可以指定 $\delta_2 > 0$, 使在任何区间 Δ 上, 只要它的长小于 δ_2 且全部被包含在上面所说的使 f 连续的一个区间内, 就有 $\omega(f;\Delta) < \dfrac{\varepsilon}{2(b-a)}$. 现在取 $\delta = \min\{\delta_1, \delta_2\}$.

设 P 是闭区间 $[a,b]$ 的任意一个分划且 $\lambda(P) < \delta$. 与分划 P 相应的和 $\sum\limits_{i=1}^{n}\omega(f;\Delta_i)\Delta x_i$ 分成两部分:

$$\Sigma\omega(f;\Delta_i)\Delta x_i = \Sigma'\omega(f;\Delta_i)\Delta x_i + \Sigma''\omega(f;\Delta_i)\Delta x_i.$$

在和 Σ' 中包含了那样的项, 与它相应的分划 P 的区间 Δ_i 与所做的间断点的 δ_1-邻域没有公共点. 对于这样的区间 Δ_i, 有 $\omega(f;\Delta_i) < \dfrac{\varepsilon}{2(b-a)}$, 因此

$$\Sigma'\omega(f;\Delta_i)\Delta x_i < \frac{\varepsilon}{2(b-a)}\Sigma'\Delta x_i \leqslant \frac{\varepsilon}{2(b-a)}(b-a) = \frac{\varepsilon}{2}.$$

容易看出, 剩下来的分划 P 的区间的长度和不超过 $(\delta+2\delta_1+\delta)k < 4\dfrac{\varepsilon}{8C\cdot k}\cdot k = \dfrac{\varepsilon}{2C}$. 因此

$$\Sigma''\omega(f;\Delta_i)\Delta x_i < C\Sigma''\Delta x_i < C\cdot\frac{\varepsilon}{2C} = \frac{\varepsilon}{2}.$$

于是, 当 $\lambda(P) < \delta$ 时, 得到

$$\sum_{i=1}^{n}\omega(f;\Delta_i)\Delta x_i < \varepsilon.$$

亦即, f 满足可积性的充分条件, 从而 $f \in \mathcal{R}[a,b]$. ►

推论 3 闭区间上的单调函数在该区间上可积.

◄ 从函数 f 在闭区间 $[a,b]$ 上的单调性推出

$$\omega(f;[a,b]) = |f(b) - f(a)|.$$

设给定了 $\varepsilon > 0$, 取 $\delta = \dfrac{\varepsilon}{|f(b)-f(a)|}$. 我们假定 $f(b) - f(a) \neq 0$, 因为在相反的情形 f 是常数, 从而无疑是可积的. 设 P 是区间 $[a,b]$ 的任意分划, 且其参数 $\lambda(P) < \delta$.

注意到 f 的单调性, 我们有

$$\begin{aligned}\sum_{i=1}^{n}\ \omega(f;\Delta_i)\Delta x_i &< \delta\sum_{i=1}^{n}\omega(f;\Delta_i) = \delta\sum_{i=1}^{n}|f(x_i) - f(x_{i-1})| \\ &= \delta\left|\sum_{i=1}^{n}(f(x_i) - f(x_{i-1}))\right| = \delta|f(b) - f(a)| = \varepsilon.\end{aligned}$$

这样, f 满足可积性的充分条件, 即 $f \in \mathcal{R}[a,b]$. ►

单调函数可能是有无穷多个 (可数多个) 间断点的函数. 例如, 在闭区间 $[0,1]$ 上由关系

$$f(x)=\begin{cases}1-\dfrac{1}{2^n}, & \text{当} 1-\dfrac{1}{2^n}\leqslant x<1-\dfrac{1}{2^{n+1}}, n\in\mathbb{N},\\ 1, & \text{当} x=1\end{cases}$$

定义的函数是不减的, 而且在每个形如 $\dfrac{1}{2^n}(n\in\mathbb{N})$ 的点处间断.

注 我们指出, 虽然我们讨论的是在闭区间上定义的实函数, 但是无论是在积分的定义中, 还是在上述那些断言的证明中, 除去推论 3, 实际上并没用到函数是实值而不是复值或向量值这样的假设条件.

相反地, 我们将要讨论的上积分和与下积分和的概念, 则是专对实值函数而言的.

定义 6 设 $f:[a,b]\to\mathbb{R}$ 是定义在闭区间 $[a,b]$ 上的有界实值函数; P 是区间 $[a,b]$ 的一个分划; $\Delta_i(i=1,\cdots,n)$ 是分划 P 的区间. 设

$$m_i=\inf_{x\in\Delta_i}f(x),\quad M_i=\sup_{x\in\Delta_i}f(x)(i=1,\cdots,n).$$

和式

$$s(f;P)=\sum_{i=1}^{n}m_i\Delta x_i$$

以及

$$S(f;P)=\sum_{i=1}^{n}M_i\Delta x_i$$

分别叫做函数 f 在区间 $[a,b]$ 上对应于分划 P 的下积分和与上积分和①.和式 $s(f;P),S(f;P)$ 也叫做关于区间 $[a,b]$ 的分划 P 的下达布和与上达布和.

如果 (P,ξ) 是闭区间 $[a,b]$ 的任一带标志点的分划, 显然有

$$s(f;P\leqslant\sigma(f;P,\xi))\leqslant S(f;P).\tag{7}$$

引理 1

$$s(f;P)=\inf_{\xi}\sigma(f;P,\xi),$$
$$S(f;P)=\sup_{\xi}\sigma(f;P,\xi).$$

①从形式上看, 这里使用 "积分和" 这一术语不太合理, 因为 m_i 和 M_i 并不总是函数 f 在某个点 $\xi_i\in\Delta_i$ 的值.

◀ 我们来验证, 例如, 对应于闭区间 $[a,b]$ 的分划 P 的上达布和, 是对应于闭区间 $[a,b]$ 的带标志点的分划 (P,ξ) 的积分和之值的上确界. 这里的上确界是关于一切可能的标志点组 $\xi=(\xi_1,\cdots,\xi_n)$ 取的.

由 (7) 可见, 只要证明: 对于任何 $\varepsilon>0$, 可以找到标志点组 $\bar{\xi}$, 使成立不等式

$$S(f;P)<\sigma(f;P,\bar{\xi})+\varepsilon. \tag{8}$$

根据数 M_i 的定义, 对每个 $i\in\{1,\cdots,n\}$ 可以找到点 $\bar{\xi}_i\in\Delta_i$ 使 $M_i<f(\bar{\xi}_i)+\dfrac{\varepsilon}{b-a}$. 设 $\bar{\xi}=(\bar{\xi}_1,\cdots,\bar{\xi}_n)$. 那么

$$\sum_{i=1}^{n}M_i\Delta x_i<\sum_{i=1}^{n}\left(f(\bar{\xi}_i)+\frac{\varepsilon}{b-a}\right)\Delta x_i=\sum_{i=1}^{n}f(\bar{\xi}_i)\Delta x_i+\varepsilon.$$

这就证明了引理的第二部分. 它的第一部分的证明是类似的. ▶

从引理的证明和不等式 (7), 考虑到黎曼积分的定义, 可以推断出下面命题是正确的.

命题 3 *有界实值函数 $f:[a,b]\to\mathbb{R}$ 在闭区间 $[a,b]$ 上黎曼可积的充分必要条件是极限*

$$\underline{I}=\lim_{\lambda(P)\to 0}s(f;P),\quad \bar{I}=\lim_{\lambda(P)\to 0}S(f;P) \tag{9}$$

存在且相等. 这时, 它们的公共值 $I=\underline{I}=\bar{I}$ 等于积分

$$\int_a^b f(x)dx.$$

◀ 事实上, 如果 (9) 式中两个极限存在并相等, 则根据极限的性质, 从 (7) 式得到积分和的极限的存在性, 而且

$$\underline{I}=\lim_{\lambda(P)\to 0}\sigma(f;P,\xi)=\bar{I}.$$

另一方面, 如果 $f\in\mathcal{R}[a,b]$, 亦即存在极限 $\lim\limits_{\lambda(P)\to 0}\sigma(f;P,\xi)=I$, 则从 (7) 和 (8) 推出存在极限

$$\lim_{\lambda(P)\to 0}S(f;P)=\bar{I}\text{且}\bar{I}=I.$$

类似地, 可以验证

$$\lim_{\lambda(P)\to 0}s(f;P)=\underline{I}=I. ▶$$

作为命题 3 的推论, 我们得到以下命题, 它是命题 2 的精确化.

命题 2′ 为使闭区间 $[a,b]$ 上的有界实值函数 $f:[a,b]\to\mathbb{R}$ 在该区间上黎曼可积, 必须且只需满足关系

$$\lim_{\lambda(P)\to 0}\sum_{i=1}^{n}\omega(f;\Delta_i)\Delta x_i=0. \tag{10}$$

◀ 考虑到命题 2, 只要证明关系 (10) 是 f 可积的必要条件.

我们知道, $\omega(f;\Delta_i)=M_i-m_i$, 因此

$$\sum_{i=1}^{n}\omega(f;\Delta_i)\Delta x_i=\sum_{i=1}^{n}(M_i-m_i)\Delta x_i=S(f;P)-s(f;P)$$

从而由命题 3 立刻从 (10) 导出 $f\in\mathcal{R}[a,b]$. ▶

c. 向量空间 $\mathcal{R}[a,b]$ 对可积函数可以进行一系列运算, 其结果不超出 $\mathcal{R}[a,b]$.

命题 4 如果 $f,g\in\mathcal{R}[a,b]$, 则

a) $(f+g)\in\mathcal{R}[a,b]$;

b) $(\alpha f)\in\mathcal{R}[a,b]$, 其中 α 是数因子;

c) $|f|\in\mathcal{R}[a,b]$;

d) $f|_{[c,d]}\in\mathcal{R}[c,d]$, 如果 $[c,d]\subset[a,b]$;

e) $(f\cdot g)\in\mathcal{R}[a,b]$.

现在我们只研究实值函数, 但性质 a),b),c),d) 对于复值函数和向量值函数也是对的. 一般说来, 对向量值函数, 乘积 $f\cdot g$ 是不定义的, 因此对它们不研究性质 e). 但这个性质对复值函数仍然有效.

现在我们在 f 和 g 是实值函数的情形下证明命题 4.

◀ a) 这个结论是显然的. 因为

$$\sum_{i=1}^{n}(f+g)(\xi_i)\Delta x_i=\sum_{i=1}^{n}f(\xi_i)\Delta x_i+\sum_{i=1}^{n}g(\xi_i)\Delta x_i$$

b) 这个结论是显然的, 因为

$$\sum_{i=1}^{n}(\alpha f)(\xi_i)\Delta x_i=\alpha\sum_{i=1}^{n}f(\xi_i)\Delta x_i.$$

c) 因为 $\omega(|f|;E)\leqslant\omega(f;E)$, 所以

$$\sum_{i=1}^{n}\omega(|f|;\Delta_i)\Delta x_i\leqslant\sum_{i=1}^{n}\omega(f;\Delta_i)\Delta x_i,$$

从而根据命题 2 得到:$(f\in\mathcal{R}[a,b])\Rightarrow(|f|\in\mathcal{R}[a,b])$.

d) 我们要证明闭区间 $[a,b]$ 上的可积函数在任何闭区间 $[c,d]\subset[a,b]$ 上的限制 $f|_{[c,d]}$ 是 $[c,d]$ 上的可积函数. 设 π 是区间 $[c,d]$ 的一个分划. 给 π 补充一些点使之成为区间 $[a,b]$ 的一个分划 P 且 $\lambda(P)\leqslant\lambda(\pi)$. 显然, 这总是可以办到的.

现在可以得到

$$\Sigma_\pi\omega(f|_{[c,d]};\Delta_i)\Delta x_i\leqslant\Sigma_P\omega(f;\Delta_i)\Delta x_i,$$

其中 Σ_π 表示关于分划 π 的全部区间取和, 而 Σ_P 表示关于分划 P 的全部区间取和.

当 $\lambda(\pi)\to 0$ 时, 根据 P 的构造方法也有 $\lambda(P)\to 0$, 于是, 根据命题 2′, 从上边的不等式得到: 如果 $[c,d]\subset[a,b]$, 则

$$(f\in\mathcal{R}[a,b])\Rightarrow(f|_{[c,d]}\in\mathcal{R}[c,d]).$$

e) 首先验证: 如果 $f\in\mathcal{R}[a,b]$, 则 $f^2\in\mathcal{R}[a,b]$.

如果 $f\in\mathcal{R}[a,b]$, 则 f 在 $[a,b]$ 上有界. 设在 $[a,b]$ 上 $|f(x)|\leqslant C<\infty$. 则

$$\begin{aligned}|f^2(x_1)-f^2(x_2)|&=|(f(x_1)+f(x_2))(f(x_1)-f(x_2))|\\&\leqslant 2C|f(x_1)-f(x_2)|,\end{aligned}$$

因此 $\omega(f^2;E)\leqslant 2C\omega(f;E)$, 其中 $E\subset[a,b]$. 这样,

$$\sum_{i=1}^{n}\omega(f^2;\Delta_i)\Delta x_i\leqslant 2C\sum_{i=1}^{n}\omega(f;\Delta_i)\Delta x_i,$$

从而根据命题 2′, 我们得到

$$(f\in\mathcal{R}[a,b])\Rightarrow(f^2\in\mathcal{R}[a,b]).$$

现在来证一般情形. 利用等式

$$(f\cdot g)(x)=\frac{1}{4}[(f+g)^2(x)-(f-g)^2(x)],$$

并注意到刚刚证明的结果和命题 4 的 a) 和 b), 就得出:

$$(f\in\mathcal{R}[a,b])\wedge(g\in\mathcal{R}[a,b])\Rightarrow(f\cdot g\in\mathcal{R}[a,b]).\ ►$$

向量空间的概念是大家从代数教程中早已知道的. 定义在一个集合上的实值函数, 可以逐点相加以及与实数相乘, 且所得到的结果仍然是定义在这个集合上的实值函数. 如果把函数看作向量, 那么可以验证, 实数域上向量空间的全部公理都成立, 从而实函数的上述集合关于函数的逐点加法以及与实数的乘法是一个向量空间.

命题 4 的 a)、b) 两点说明了可积函数相加以及与数相乘的结果均不超出 $\mathcal{R}[a,b]$. 这样一来, $\mathcal{R}[a,b]$ 本身也是一个向量空间, 它是定义在闭区间 $[a,b]$ 上的实值函数所构成的向量空间的子空间.

d. 函数黎曼可积性的勒贝格判别法 最后, 我们暂不加证明地给出勒贝格定理, 它刻画了黎曼可积函数的内在本质.

为此我们引进以下概念, 而它本身也是有用的.

定义 7 称集合 $E \subset \mathbb{R}$ (在勒贝格意义下) *有零测度或称它是一个零测度集*, 如果对于任何 $\varepsilon > 0$, 存在集合 E 的由最多可数个开区间组成的覆盖 $\{I_k\}$, 且这些区间的长度和 $\sum\limits_{k=1}^{\infty} |I_k|$ 不超过 ε.

由于级数 $\sum\limits_{k=1}^{\infty} |I_k|$ 是绝对收敛的, 求和的顺序对结果没有影响 (见第五章,§5.2 的命题 4), 因此上面的定义是合理的.

引理 2 a) *一个点或有限多个点的集合是零测度集.*

b) *有限多个或可数多个零测度集的并是零测度集.*

c) *零测度集的子集本身也是零测度集.*

d) *当 $a < b$ 时, 区间 $[a, b]$ 不是零测度集.*

◀ a) 点可以用一个长度小于事先指定的任意小的 $\varepsilon > 0$ 的开区间覆盖, 因此点是零测度集. a) 的其余部分可由 b) 推出.

b) 设 $E = \bigcup\limits_n E^n$ 是最多可数个零测度集E^n 的并. 对于每个 E^n, 关于 $\varepsilon > 0$, 做 E^n 的覆盖 $\{I_k^n\}$ 使 $\sum\limits_k |I_k^n| < \varepsilon/2^n$.

由于有限或可数集的有限或可数并本身最多是可数的, 开区间 $I_k^n, k, n \in \mathbb{N}$, 所组成的 E 的覆盖最多也只能是可数的, 而且

$$\sum_{n,k} |I_k^n| \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2^2} + \cdots + \frac{\varepsilon}{2^n} + \cdots = \varepsilon.$$

在 $\sum\limits_{n,k} |I_k^n|$ 中关于指标 n 和 k 的求和顺序是无关紧要的, 因为该级数只要在某求和顺序下收敛于一个和, 那么它在任何求和顺序下收敛于同一个和. 由于我们的级数的部分和以 ε 为上界, 因此级数在某一求和顺序下的收敛性是显然的.

于是, E 是勒贝格意义下的零测度集.

c) 这个断言显然可由零测度集的定义和覆盖的定义直接推出.

d) 由于闭区间的任一个由开区间作成的覆盖有有限覆盖, 它的开区间的长度和显然不超过原覆盖的开区间的长度和. 因此, 只要证明组成区间 $[a, b]$ 的任意一个有限覆盖的开区间的长度和不小于这个区间的长度 $b - a$ 即可.

我们关于覆盖的开区间数进行归纳证明.

当 $n = 1$ 时, 也就是区间 $[a, b]$ 被一个开区间 (α, β) 包含的情形, 显然 $\alpha < a < b < \beta$, 从而 $\beta - \alpha > b - a$.

假设对于 $n = 1, \cdots, k$ 我们的断言正确, 现在考察覆盖是由 $k+1$ 个开区间组成的情形. 取开区间 (α_1, α_2), 它覆盖点 a. 如果 $\alpha_2 \geqslant b$, 则 $\alpha_2 - \alpha_1 > b - a$, 从而证明

了我们的断言. 如果 $a < \alpha_2 < b$, 则区间 $[\alpha_2, b]$ 就由最多 k 个区间所组成的开区间系所覆盖, 根据归纳假定, 它的区间长度的和不小于 $b - \alpha_2$. 但是

$$b - a = (b - \alpha_2) + (\alpha_2 - a) < (b - \alpha_2) + (\alpha_2 - \alpha_1),$$

从而, 区间 $[a, b]$ 原来的覆盖所含开区间的长度和大于 $b - a$. ▶

有趣的是, 根据引理 2 的 a) 和 b), 数轴 $\mathbb{R}$ 上的全部有理点的集合 $\mathbb{Q}$ 是一个零测度集. 当与该引理的 a) 比较时, 乍一看, 这似乎是不可想象的.

定义 8 在集合 X 上, 如果除去零测度集合的点, 某一性质都成立, 则说该性质在集合 X 上*几乎处处* 成立, 或者说在集合 X 的*几乎所有点*有该性质.

现在来叙述勒贝格定理 (判别法).

定理 *定义在闭区间 $[a, b]$ 上的函数 f, 当且仅当它在 $[a, b]$ 上有界且几乎处处连续时, 它在该区间上黎曼可积*, 即

$$(f \in \mathcal{R}[a, b]) \Leftrightarrow (f\text{在}[a, b]\text{上有界}) \wedge (f\text{在}[a, b]\text{上几乎处处连续}).$$

显然, 命题 2 的推论 1,2,3 和命题 4 都可以从勒贝格判别法和引理 2 中所说的零测度集的性质推出.

现在我们不去证明勒贝格判别法, 因为我们面临的任务是研究那些足够正规的函数, 暂时还不需要这个准则. 但是, 现在已经可以完全清楚地阐明勒贝格准则的思想内容了.

命题 2′ 中的关系 (10) 包含了实值函数可积性判别准则. 为使和 $\sum\limits_{i=1}^{n} \omega(f; \Delta_i)\Delta x_i$ 很小, 首先考虑因子 $\omega(f; \Delta_i)$. 我们知道, 它在函数的连续点的小邻域中是很小的. 如果 $\{\Delta_i\}$ 中有些区间含有函数的间断点, 那么, 对于这些区间, 不管区间 $[a, b]$ 的分划 P 多么细, $\omega(f; \Delta_i)$ 都不接近于零. 但是, $\omega(f; \Delta_i) \leqslant \omega(f; [a, b]) < \infty$ (因为 f 在 $[a, b]$ 上有界). 因此, 含有间断点的那些项的和还是可能很小的, 只要分划中覆盖间断点集合的那些区间的长度和是很小的, 或者更准确些说, 只要函数在这些区间上振幅的高度可以从这些区间的长度很小得到补偿.

勒贝格判别准则正是这些直观东西的严格叙述和实现.

现在引入用以说明黎曼可积函数性质的两个经典例子.

例 1 狄利克雷函数.

$$\mathcal{D}(x) = \begin{cases} 1, \text{当}x \in \mathbb{Q}, \\ 0, \text{当}x \in \mathbb{R}\backslash\mathbb{Q} \end{cases}$$

在闭区间 $[0,1]$ 上不可积, 因为对区间 $[0,1]$ 的任何分划 P, 在其每个区间 Δ_i 中都能找到一个有理数 ξ_i' 和一个无理数 ξ_i''. 因此

$$\sigma(f;P,\xi')=\sum_{i=1}^{n}1\cdot\Delta x_i=1,$$

而

$$\sigma(f;P,\xi'')=\sum_{i=1}^{n}0\cdot\Delta x_i=0.$$

这样一来, 当 $\lambda(P)\to 0$ 时函数 $\mathcal{D}(x)$ 的积分和不可能有极限.

从勒贝格判别法来看, 狄利克雷函数的不可积性也是明显的, 因为函数 $\mathcal{D}(x)$ 在区间 $[0,1]$ 中间断点的集合不是零测度集 (见引理 2).

例 2 黎曼函数

$$\mathcal{R}(x)=\begin{cases}\dfrac{1}{n}, & 当x\in\mathbb{Q}且x=\dfrac{m}{n}是既约分数,\\ 0, & 当x\in\mathbb{R}\backslash\mathbb{Q}.\end{cases}$$

我们已经在第四章 §2 的第 2 段中研究过这个函数, 而且知道, 函数 $\mathcal{R}(x)$ 在一切无理点连续, 在一切有理点间断, 这样一来, 函数 $\mathcal{R}(x)$ 的间断点集合是可数的, 从而其测度为零, 根据勒贝格判别法, 尽管这个函数在积分区间的任何分划的任何区间中都有间断点, 但它在任何区间 $[a,b]\subset\mathbb{R}$ 上仍然是可积的.

例 3 现在再考察一个不太经典的问题和例子.

设 $f:[a,b]\to\mathbb{R}$ 是 $[a,b]$ 区间上的可积函数, 它在区间 $[c,d]$ 中取值. 又设函数 $g:[c,d]\to\mathbb{R}$ 是连续的. 因此, 复合函数 $g\circ f:[a,b]\to\mathbb{R}$ 在整个区间 $[a,b]$ 上定义并在使 f 连续的那些点上连续. 由此, 根据勒贝格判别准则得到, $(g\circ f)\in\mathcal{R}[a,b]$.

现在举例说明, 可积函数的复合函数未必可积.

考察函数 $g(x)=|\text{sgn}|(x)$, 当 $x\neq 0$ 时这个函数取值为 1; 当 $x=0$ 时, 它取值为零. 设 f 是上面所讲的黎曼函数. 可直接验证, 譬如在区间 $[1,2]$ 上, 复合函数 $(g\circ f)(x)$ 正好是狄利克雷函数$\mathcal{D}(x)$. 这样一来, 甚至 $g(x)$ 只有一个间断点, 也可导致复合函数 $g\circ f$ 不可积.

练 习

1. 达布定理

 a) 设 $s(f;P)$ 和 $S(f;P)$ 是在闭区间 $[a,b]$ 上定义的有界实值函数对应于这个区间的分划 P 的下达布和及上达布和. 试证: 对于闭区间 $[a,b]$ 的任两个分划 P_1,P_2, 成立

$$s(f;P_1)\leqslant S(f;P_2).$$

b) 设分划 $\tilde{P}$ 是闭区间 $[a,b]$ 的分划 P 的开拓, 而 $\Delta_{i_1},\cdots,\Delta_{i_k}$ 是分划 P 的那样一些区间, 它们包含着属于分划 $\tilde{P}$ 但不属于分划 P 的点. 试证: 成立如下估计:

$$0\leqslant S(f;P)-S(f;\tilde{P})\leqslant \omega(f;[a,b])(\Delta x_{i_1}+\cdots+\Delta x_{i_k}),$$
$$0\leqslant s(f;\tilde{P})-s(f;P)\leqslant \omega(f;[a,b])(\Delta x_{i_1}+\cdots+\Delta x_{i_k}).$$

c) $\underline{I}=\sup\limits_P s(f;P),\overline{I}=\inf\limits_P S(f;P)$ 分别叫做函数 f 在闭区间 $[a,b]$ 上的下达布积分及上达布积分. 试证: $\underline{I}\leqslant\overline{I}$.

d) 试证达布定理

$$\underline{I}=\lim_{\lambda(P)\to 0}s(f;P);\overline{I}=\lim_{\lambda(P)\to 0}S(f;P).$$

e) 试证:$(f\in\mathcal{R}[a,b])\Leftrightarrow(\underline{I}=\overline{I})$.(达布判别法)

f) 试证:$f\in\mathcal{R}[a,b]$, 当且仅当对于任何 $\varepsilon>0$ 存在区间 $[a,b]$ 的分划 P 使 $S(f;P)-s(f;P)<\varepsilon$.

2. 康托尔零测度集

a) 第二章 §4 的问题 7 中所说的康托尔集是不可数的. 试验证, 它仍是勒贝格意义下的零测度集. 试问: 怎样修改康托尔集的构造方法可以得到类似的处处"有窟窿"但测度不是零的集合.(这种集合也叫康托尔集).

b) 设在区间 $[0,1]$ 上给定了一个函数, 它在某康托尔集外等于零, 而在该康托尔集上等于 1. 试证: 当且仅当该康托尔集是零测度集时, 给定的函数是黎曼可积的.

c) 在区间 $[0,1]$ 上构造单调不减、连续的非常数函数, 而且最多除去一个具零测度的康托尔集, 它的导数处处等于 0.

3. 勒贝格判别法

a) 直接验证 (不准引用勒贝格判别法) 本节例 2 中的黎曼函数的可积性.

b) 试证: 有界函数 $f\in\mathcal{R}[a,b]$, 当且仅当对于任意的 $\varepsilon>0$ 和 $\delta>0$, 存在区间 $[a,b]$ 的一个分划 P, 使函数在其上的振幅大于 ε 的那些区间的长度之和不超过 δ.

c) 试证:$f\in\mathcal{R}[a,b]$, 当且仅当 f 在 $[a,b]$ 上有界且对于任何 $\varepsilon>0$ 和 $\delta>0$, 区间 $[a,b]$ 上使 f 的振幅大于 ε 的点的集合可以用有限多个开区间覆盖, 而这些开区间的长度和小于 δ.(杜布瓦雷蒙[①]判别法)

d) 利用上题结果证明函数黎曼可积性的勒贝格判别法.

4. 试证: 如果 $f,g\in\mathcal{R}[a,b]$ 且 f,g 是实的, 则 $\max\{f,g\}\in\mathcal{R}[a,b],\min\{f,g\}\in\mathcal{R}[a,b]$.

5. 试证:

a) 如果 $f,g\in\mathcal{R}[a,b]$, 在 $[a,b]$ 上几乎处处有 $f(x)=g(x)$, 则 $\int_a^b f(x)dx=\int_a^b g(x)dx$.

b) 如果 $f\in\mathcal{R}[a,b]$, 在 $[a,b]$ 上几乎处处有 $f(x)=g(x)$, 则甚至当 g 在 $[a,b]$ 上定义且有界时, 它也可能是黎曼不可积的.

6. 向量值函数的积分

a) 设 $\boldsymbol{r}(t)$ 是空间中一动点的矢径向量; $\boldsymbol{r}_0=\boldsymbol{r}(0)$ 是该点的初始位置; $\boldsymbol{v}(t)$ 是作为时间的函数的速度向量. 试根据 $\boldsymbol{r}_0$ 和函数 $\boldsymbol{v}(t)$ 求 $\boldsymbol{r}(t)$.

[①] 杜布瓦雷蒙(du Bois Raymond) (1831—1889)—— 德国数学家.

b) 能否把向量值函数的积分归结为实值函数的积分?

c) 在命题 2′ 中的可积性判别法对于向量值函数是否成立?

d) 对于向量值函数, 勒贝格判别法是否成立?

e) 本节的哪些概念和结果可以转移到复值函数的情形?

§2. 积分的线性性、可加性和单调性

1. 作为空间 $\mathcal{R}[a,b]$ 上的线性函数的积分

定理 1 *如果 f 和 g 是闭区间 $[a,b]$ 上的可积函数, 则它们的线性组合 $\alpha f+\beta g$ 也是 $[a,b]$ 上的可积函数, 且*

$$\int_a^b(\alpha f+\beta g)(x)dx=\alpha\int_a^b f(x)dx+\beta\int_a^b g(x)dx. \tag{1}$$

◀ 考虑关系 (1) 左端积分的积分和并对它进行变换, 有

$$\sum_{i=1}^n(\alpha f+\beta g)(\xi_i)\Delta x_i=\alpha\sum_{i=1}^n f(\xi_i)\Delta x_i+\beta\sum_{i=1}^n g(\xi_i)\Delta x_i. \tag{2}$$

因为当分划的参数 $\lambda(P)$ 趋于 0 时,(2) 的右端趋于 (1) 右边的积分的线性组合, 所以 (2) 的左边当 $\lambda(p)\to 0$ 时也有极限, 而且这个极限与右边的极限相等. 这样, $(\alpha f+\beta g)\in\mathcal{R}[a,b]$ 且等式 (1) 成立. ▶

如果把集合 $\mathcal{R}[a,b]$ 看做是实数域上的向量空间, 而积分 $\int_a^b f(x)dx$ 看做是定义在向量空间 $\mathcal{R}[a,b]$ 上的实值函数, 则定理 1 说明, 积分是向量空间 $\mathcal{R}[a,b]$ 上的线性函数.

为避免混淆, 函数的函数通常叫做*泛函*. 于是, 我们证明了: 积分是可积函数向量空间上的线性泛函.

2. 作为积分区间的可加函数的积分 积分值 $\int_a^b f(x)dx=I(f;[a,b])$, 既依赖于被积函数, 又依赖于进行积分的区间, 譬如, 如果 $f\in\mathcal{R}[a,b]$, 则对于 $[\alpha,\beta]\subset[a,b]$ 有 $f|_{[\alpha,\beta]}\in\mathcal{R}[\alpha,\beta]$, 亦即积分 $\int_\alpha^\beta f(x)dx$ 有定义, 从而可以用它对积分区间 $[\alpha,\beta]$ 的依赖关系的观点研究它.

引理 1 *如果 $a<b<c, f\in\mathcal{R}[a,c]$, 则 $f|_{[a,b]}\in\mathcal{R}[a,b], f|_{[b,c]}\in\mathcal{R}[b,c]$ 且* [①]

$$\int_a^c f(x)dx=\int_a^b f(x)dx+\int_b^c f(x)dx. \tag{3}$$

①我们提醒注意, 符号 $f|_E$ 表示函数 f 在集合 E 上的限制, 其中 E 被包含在 f 的定义域内. 在等式 (3) 的右边本应写 f 在相应的区间上的限制, 而不应当写成 f.

◀ 首先注意, 有界函数 f 在区间 $[a,b]$ 和 $[b,c]$ 上的可积性由上一节命题 4 保证. 其次, 由于 $f \in \mathcal{R}[a,b]$, 所以当把积分 $\int_a^c f(x)dx$ 作为积分和的极限进行计算时, 我们可以选取任意的对我们方便的区间 $[a,c]$ 的分划, 我们只取区间 $[a,c]$ 的那样一些分划, 它们都包含点 b. 显然, 每一个这样的带标志点的分划 (P,ξ) 都产生区间 $[a,b]$ 和 $[b,c]$ 的相应的分划 (P',ξ') 和 (P'',ξ''), 而且 $P = P' \cup P''$ 以及 $\xi = \xi' \cup \xi''$. 由此得到积分和之间的关系:

$$\sigma(f;P,\xi) = \sigma(f;P',\xi') + \sigma(f;P'',\xi'').$$

因为 $\lambda(P') \leqslant \lambda(P)$ 和 $\lambda(P'') \leqslant \lambda(P)$, 所以当 $\lambda(P)$ 充分小时, 上面的每个积分和就充分接近于 (3) 中相应的积分. 这样就证明了 (3) 成立. ▶

为了对所得的结果的适用性做一些推广, 暂时回到积分定义上去.

我们已经把积分定义为与积分区间 $[a,b]$ 的带标志点的分划 (P,ξ) 相应的积分和

$$\sigma(f;P,\xi) = \sum_{i=1}^{n} f(\xi_i)\Delta x_i \tag{4}$$

的极限, 分划 P 由一个单调的有限点列 $x_0, x_1, \cdots, x_n$ 组成, 而且点 x_0 与积分下限 a 重合, 点 x_n 与积分上限 b 重合. 这种构造当时是在 $a < b$ 的假定下进行的. 如果现在取任意两个数 a, b, 不要求 $a < b$, 把 a 作为积分下限, 而 b 作为上限, 进行上面那种构造, 那么我们将重新得到形如 (4) 的和; 当 $a < b$ 时有 $\Delta x_i > 0 (i = 1, \cdots, n)$, 而当 $a > b$ 时有 $\Delta x_i < 0 (i = 1, \cdots, n)$, 因为 $\Delta x_i = x_i - x_{i-1}$. 这样一来, 当 $a < b$ 时的和 (4) 与区间 $[b,a] (b < a)$ 相应分划的积分和只差一个符号.

根据这些想法, 当 $a > b$ 时, 我们做如下约定:

$$\int_a^b f(x)dx := -\int_b^a f(x)dx. \tag{5}$$

与此同时, 自然地也假定

$$\int_a^a f(x)dx := 0. \tag{6}$$

做了这些约定之后, 并注意到引理 1, 我们将得到积分的如下重要性质.

定理 2 *设 $a, b, c \in \mathbb{R}$, 而 f 是在以这些点为端点的最大区间上的可积函数, 则 f 在另外两个区间上的限制在相应的区间上也可积, 且成立以下等式*

$$\int_a^b f(x)dx + \int_b^c f(x)dx + \int_c^a f(x)dx = 0. \tag{7}$$

◀ 由于等式 (7) 关于 a, b, c 是对称的, 不失一般性可以假定 $a = \min\{a,b,c\}$, 如果 $\max\{a,b,c\} = c$ 和 $a < b < c$, 则根据引理 1 有

$$\int_a^b f(x)dx + \int_b^c f(x)dx - \int_a^c f(x)dx = 0.$$

由此, 注意到约定 (5), 就得到等式 (7).

如果 $\max\{a,b,c\}=b$ 且 $a<c<b$, 则根据引理 1 有

$$\int_a^c f(x)dx+\int_c^b f(x)dx-\int_a^b f(x)dx=0.$$

由此, 注意到约定 (5), 也得到等式 (7).

最后, 如果点 a,b,c 中有两个或三个重合在一起, 则 (7) 可由约定 (5) 和 (6) 直接得到. ►

定义1 设每个由闭区间 $[a,b]$ 的点 α,β 组成的有序点对 (α,β) 对应于一个数 $I(\alpha,\beta)$, 而且对于任意三点 $\alpha,\beta,\gamma\in[a,b]$, 成立等式

$$I(\alpha,\gamma)=I(\alpha,\beta)+I(\beta,\gamma).$$

那么, 函数 $I(\alpha,\beta)$ 就叫做在包含于闭区间 $[a,b]$ 中的区间上定义的定向区间的可加函数.

如果 $f\in\mathcal{R}[A,B]$, 而 $a,b,c\in[A,B]$, 令

$$I(a,b)=\int_a^b f(x)dx,$$

则由 (7) 得

$$\int_a^c f(x)dx=\int_a^b f(x)dx+\int_b^c f(x)dx, \tag{8}$$

也就是说, 积分是积分区间的可加函数, 在这种情形, 区间的定向性在于我们给积分区间的端点赋了顺序, 指出了哪是第一个 (即积分下限), 哪是第二个 (即积分上限).

3. 积分的估计, 积分的单调性和中值定理

a. 积分的一个一般估计 我们从积分的一个一般估计开始, 以后会看到, 它的正确性不限于实值函数的积分.

定理 3 *如果 $a\leqslant b, f\in\mathcal{R}[a,b]$, 则 $|f|\in\mathcal{R}[a,b]$ 且成立*

$$\left|\int_a^b f(x)dx\right|\leqslant\int_a^b |f|(x)dx. \tag{9}$$

如果在 $[a,b]$ 上还假定 $|f|(x)\leqslant C$, 那么

$$\int_a^b |f|(x)dx\leqslant C(b-a). \tag{10}$$

◄ 当 $a=b$ 时, 定理的断言是显然的, 因此, 我们假定 $a<b$.

为了证明这个定理, 只要记起 $|f| \in \mathcal{R}[a,b]$(见 §1 的命题 4), 就得到积分和 $\sigma(f;P,\xi)$ 的以下估计:

$$\left|\sum_{i=1}^{n} f(\xi_i)\Delta x_i\right| \leqslant \sum_{i=1}^{n} |f(\xi_i)||\Delta x_i| = \sum_{i=1}^{n} |f(\xi_i)|\Delta x_i \leqslant C\sum_{i=1}^{n}\Delta x_i = C(b-a).$$

令 $\lambda(P) \to 0$ 取极限, 我们得到

$$\left|\int_a^b f(x)dx\right| \leqslant \int_a^b |f(x)|dx \leqslant C(b-a). ►$$

b. 积分的单调性和第一中值定理 下面的定理是实值函数的积分所特有的.

定理 4 *如果 $a \leqslant b, f_1$、$f_2 \in \mathcal{R}[a,b]$ 且 $f_1(x) \leqslant f_2(x)$ 在每一点 $x \in [a,b]$ 都成立, 则*

$$\int_a^b f_1(x)dx \leqslant \int_a^b f_2(x)dx. \tag{11}$$

◄ 当 $a=b$ 时定理的结论是显然的. 如果 $a<b$, 则关于积分和成立不等式

$$\sum_{i=1}^{n} f_1(\xi_i)\Delta x_i \leqslant \sum_{i=1}^{n} f_2(\xi_i)\Delta x_i$$

(因为 $\Delta x_i > 0(i=1,\cdots,n)$). 只要在这个不等式中令 $\lambda(P) \to 0$ 取极限即可. ►

定理 4 表明积分关于被积函数是单调的.

从定理 4 可以得到一系列有用的推论.

推论 1 *如果 $a \leqslant b, f \in \mathcal{R}[a,b]$ 且 $m \leqslant f(x) \leqslant M$ 在 $x \in [a,b]$ 成立, 则*

$$m(b-a) \leqslant \int_a^b f(x)dx \leqslant M(b-a). \tag{12}$$

特别地, 如果 $0 \leqslant f(x)$ 在 $x \in [a,b]$ 成立, 则

$$0 \leqslant \int_a^b f(x)dx.$$

◄ 积分不等式 $m \leqslant f(x) \leqslant M$ 的每一项并引用定理 4 就得到关系 (12). ►

推论 2 *如果 $f \in \mathcal{R}[a,b], m = \inf\limits_{x\in[a,b]} f(x), M = \sup\limits_{x\in[a,b]} f(x)$, 则存在 $\mu \in [m,M]$ 使*

$$\int_a^b f(x)dx = \mu(b-a). \tag{13}$$

◀ 如果 $a=b$, 结论是显然的, 如果 $a \neq b$, 则令

$$\mu=\frac{1}{b-a}\int_a^b f(x)dx.$$

于是, 当 $a<b$, 从 (12) 可以推出 $m \leqslant \mu \leqslant M$. 但是, 将 (13) 两端同时改变符号, 即可看出它对 $b<a$ 的情形也成立. ▶

推论 3 *如果 $f \in C[a,b]$, 则存在 $\xi \in [a,b]$ 使*

$$\int_a^b f(x)dx = f(\xi)(b-a). \tag{14}$$

◀ 根据连续函数的中间值定理, 在区间 $[a,b]$ 上存在点 ξ 使 $f(\xi)=\mu$, 如果 μ 满足条件

$$m=\min_{x\in[a,b]} f(x) \leqslant \mu \leqslant \max_{x\in[a,b]} f(x) = M.$$

这样一来,(14) 可由 (13) 导出. ▶

等式 (14) 常叫做积分的*第一中值定理*. 对以下更一般的断言我们仍然保留这个名称.

定理 5 (积分的第一中值定理) *设 $f,\ g \in \mathcal{R}[a,\ b]$, $m=\inf\limits_{x\in[a,\ b]} f(x)$, $M=\sup\limits_{x\in[a,b]} f(x)$. 如果函数 g 在区间 $[a,b]$ 上非负 (或非正), 则*

$$\int_a^b (f\cdot g)(x)dx = \mu\int_a^b g(x)dx, \tag{15}$$

其中 $\mu \in [m,M]$.

如果还有 $f \in C[a,b]$, 则存在 $\xi \in [a,b]$ 使

$$\int_a^b (f\cdot g)(x)dx = f(\xi)\int_a^b g(x)dx. \tag{16}$$

◀ 因为在等式 (15) 中两边调换积分上、下限将导致两边同时变号, 所以只要对 $a<b$ 的情形来证明这个等式就行了. 改变 $g(x)$ 的符号也同时使 (15) 两端变号, 所以不失一般性, 可以假定 $g(x) \geqslant 0$ 在 $x \in [a,b]$ 成立.

由于 $m=\inf\limits_{x\in[a,b]} f(x)$ 和 $M=\sup\limits_{x\in[a,b]} f(x)$, 因此当 $g(x)\geqslant 0$ 时就有

$$mg(x) \leqslant f(x)g(x) \leqslant Mg(x).$$

因为 $m\cdot g \in \mathcal{R}[a,b]$, $f\cdot g \in \mathcal{R}[a,b]$ 和 $M\cdot g \in \mathcal{R}[a,b]$, 所以, 应用定理 4 和定理 1, 我们得到

$$m\int_a^b g(x)dx \leqslant \int_a^b f(x)g(x)dx \leqslant M\int_a^b g(x)dx. \tag{17}$$

如果 $\int_a^b g(x)dx = 0$, 很显然, 从这个不等式能导出关系 (15).

如果 $\int_a^b g(x)dx \neq 0$, 令 $\mu = \left(\int_a^b g(x)dx\right)^{-1} \int_a^b f(x)g(x)dx$, 则从 (17) 得到

$$m \leqslant \mu \leqslant M,$$

从而证明了关系 (15).

最后, 等式 (16) 由 (15) 和关于函数 $f \in C[a,b]$ 的中间值定理并注意, 在 $f \in C[a,b]$ 的情形, 有

$$m = \min_{x\in[a,b]} f(x) \quad 和 \quad M = \max_{x\in[a,b]} f(x)$$

即可得到.

我们指出, 如果 $g(x) \equiv 1$ 在 $[a,b]$ 上, 等式 (14) 由 (16) 得出.

c. 积分的第二中值定理 在黎曼积分理论中更加特别, 更加微妙的是第二中值定理[①].

为了不使这个定理的证明过分复杂, 首先做一些准备, 它们本身也是很有趣的.

阿贝尔变换 阿贝尔变换是和式 $\sum_{i=1}^{n} a_i b_i$ 的如下变换. 设 $A_k = \sum_{i=1}^{k} a_i$, 还令 $A_0 = 0$. 于是有

$$\sum_{i=1}^{n} a_i b_i = \sum_{i=1}^{n} (A_i - A_{i-1}) b_i = \sum_{i=1}^{n} A_i b_i - \sum_{i=1}^{n} A_{i-1} b_i$$
$$= \sum_{i=1}^{n} A_i b_i - \sum_{i=0}^{n-1} A_i b_{i+1} = A_n b_n - A_0 b_1 + \sum_{i=1}^{n-1} A_i (b_i - b_{i+1}).$$

因此,

$$\sum_{i=1}^{n} a_i b_i = (A_n b_n - A_0 b_1) + \sum_{i=1}^{n-1} A_i (b_i - b_{i+1}), \tag{18}$$

或

$$\sum_{i=1}^{n} a_i b_i = A_n b_n + \sum_{i=1}^{n-1} A_i (b_i - b_{i+1}), \tag{19}$$

因为 $A_0 = 0$.

根据阿贝尔变换容易证明

引理 2 如果数 $A_k = \sum_{i=1}^{k} a_i (k = 1, \cdots, n)$ 满足不等式 $m \leqslant A_k \leqslant M$, 而数 $b_i (i = 1, \cdots, n)$ 是非负的且 $b_i \geqslant b_{i+1} (i = 1, \cdots, n-1)$, 则

$$mb_1 \leqslant \sum_{i=1}^{n} a_i b_i \leqslant Mb_1. \tag{20}$$

[①] 对函数辅加一些通常完全可以接受的条件后, 这一段的基本定理 6 很容易从第一中值定理推出. 关于这一点, 可参看下一节的问题3.

◀ 利用 $b_n \geqslant 0$ 和 $b_i - b_{i+1} \geqslant 0 (i = 1, \cdots, n-1)$ 从 (19) 得到

$$\sum_{i=1}^{n} a_i b_i \leqslant M b_n + \sum_{i=1}^{n-1} M(b_i - b_{i+1}) = M b_n + M(b_1 - b_n) = M b_1.$$

用类似的方法也可验证 (20) 中左边的不等式. ▶

引理 3 *如果 $f \in \mathcal{R}[a,b]$, 则对于 $x \in [a,b]$ 可以定义函数*

$$F(x) = \int_a^x f(t)dt \tag{21}$$

且 $F \in C[a,b]$.

◀ 对于任何 $x \in [a,b]$, 积分 (21) 的存在性已从 §1 的命题 4 知道, 因此剩下的只是验证函数 $F(x)$ 的连续性. 由于 $f \in \mathcal{R}[a,b]$, 在 $[a,b]$ 上我们有 $|f(x)| \leqslant C < \infty$. 设 $x \in [a,b]$ 和 $x+h \in [a,b]$. 那么, 根据积分的可加性以及不等式 (9),(10), 我们得到

$$\begin{aligned} |F(x+h) - F(x)| &= \left| \int_a^{x+h} f(t)dt - \int_a^x f(t)dt \right| \\ &= \left| \int_x^{x+h} f(t)dt \right| \leqslant \left| \int_x^{x+h} |f(t)|dt \right| \leqslant C|h|. \end{aligned}$$

这里我们利用了不等式 (10), 并注意到以下事实: 当 $h < 0$ 时, 有

$$\left| \int_x^{x+h} |f(t)|dt \right| = \left| - \int_{x+h}^{x} |f(t)|dt \right| = \int_{x+h}^{x} |f(t)|dt.$$

这样, 我们证明了: 如果 $x, x+h \in [a,b]$, 则

$$|F(x+h) - F(x)| \leqslant C|h|. \tag{22}$$

显然, 由此立即导出函数 F 在区间 $[a,b]$ 的任意点连续. ▶

现在来证明下面的引理, 它是第二中值定理的变形.

引理 4 *如果 $f, g \in \mathcal{R}[a,b]$, 而 g 在区间 $[a,b]$ 上非负, 不增, 则存在 $\xi \in [a,b]$ 使*

$$\int_a^b (f \cdot g)(x)dx = g(a) \int_a^{\xi} f(x)dx. \tag{23}$$

在证明这个引理之前, 首先指出, 与第一中值定理中的关系 (16) 不同,(23) 中积分号下的函数是 f, 而不是单调函数 g.

◀ 为了证明公式 (23), 跟已经研究过的上述情形一样, 我们将设法估计相应的积分和.

设 P 是闭区间 $[a,b]$ 的一个分划. 首先记下等式

$$\int_a^b (f\cdot g)(x)dx = \sum_{i=1}^n \int_{x_{i-1}}^{x_i} (f\cdot g)(x)dx$$
$$= \sum_{i=1}^n g(x_{i-1}) \int_{x_{i-1}}^{x_i} f(x)dx$$
$$+ \sum_{i=1}^n \int_{x_{i-1}}^{x_i} [g(x) - g(x_{i-1})]f(x)dx.$$

我们来证明, 当 $\lambda(P)\to 0$ 时, 上式中最后的和式趋于 0.

因为 $f\in\mathcal{R}[a,b]$, 所以 $|f(x)|\leqslant C<\infty$ 在 $[a,b]$ 上成立. 于是, 利用证明了的积分的性质, 得到: 当 $\lambda(P)\to 0$ 时, 有

$$\left|\sum_{i=1}^n \int_{x_{i-1}}^{x_i} [g(x)-g(x_{i-1})]f(x)dx\right| \leqslant \sum_{i=1}^n \int_{x_{i-1}}^{x_i} |g(x)-g(x_{i-1})||f(x)|dx$$
$$\leqslant C\sum_{i=1}^n \int_{x_{i-1}}^{x_i} |g(x)-g(x_{i-1})|dx$$
$$\leqslant C\sum_{i=1}^n \omega(g,\Delta_i)\Delta x_i \to 0,$$

因为 $g\in\mathcal{R}[a,b]$(见 §1, 命题 2). 这样就证明了

$$\int_a^b (f\cdot g)(x)dx = \lim_{\lambda(P)\to 0} \sum_{i=1}^n g(x_{i-1}) \int_{x_{i-1}}^{x_i} f(x)dx. \tag{24}$$

现在估计 (24) 右端的和式, 令 $F(x)=\int_a^x f(t)dt$, 根据引理 3, 我们得到, 它在区间 $[a,b]$ 上连续.

令

$$m = \min_{x\in[a,b]} F(x), M = \max_{x\in[a,b]} F(x).$$

由于 $\int_{x_{i-1}}^{x_i} f(x)dx = F(x_i)-F(x_{i-1})$, 所以

$$\sum_{i=1}^n g(x_{i-1}) \int_{x_{i-1}}^{x_i} f(x)dx = \sum_{i=1}^n [F(x_i)-F(x_{i-1})]g(x_{i-1}) \tag{25}$$

考虑到函数 g 在区间 $[a,b]$ 上的非负性而且是单调不增的, 记

$$a_i = F(x_i)-F(x_{i-1}), b_i = g(x_{i-1}),$$

根据引理 2, 有

$$mg(a) \leqslant \sum_{i=1}^n [F(x_i)-F(x_{i-1})]g(x_i) \leqslant Mg(a). \tag{26}$$

因为

$$A_k = \sum_{i=1}^{k} a_i = F(x_k) - F(x_0) = F(x_k) - F(a) = F(x_k).$$

这样, 我们证明了: 和式 (25) 满足不等式 (26). 因此, 利用关系 (24), 得

$$mg(a) \leqslant \int_a^b (f \cdot g)(x)dx \leqslant Mg(a). \tag{27}$$

如果 $g(a) = 0$, 那么, 如同 (27) 的证明, 要证明的关系 (23) 显然是对的.

如果 $g(a) > 0$, 令

$$\mu = \frac{1}{g(a)} \int_a^b (f \cdot g)(x)dx.$$

从 (27) 得 $m \leqslant \mu \leqslant M$; 而由函数 $F(x) = \int_a^x f(t)dt$ 在区间 $[a, b]$ 上的连续性可知, 存在点 $\xi \in [a, b]$ 使 $F(\xi) = \mu$, 正是这个结果确定了等式 (23). ►

定理 6 (积分的第二中值定理) *如果 $f, g \in \mathcal{R}[a, b]$ 而 g 在 $[a, b]$ 上单调, 则存在点 $\xi \in [a, b]$ 使*

$$\int_a^b (f \cdot g)dx = g(a) \int_a^\xi f(x)dx + g(b) \int_\xi^b f(x)dx. \tag{28}$$

等式 (28)(以及等式 (23)) 常叫做波内公式 [①]

◄ 设 g 是 $[a, b]$ 区间上的不减函数. 则 $G(x) = g(b) - g(x)$ 是 $[a, b]$ 上的非负, 不增, 可积的函数. 利用公式 (23), 我们得到

$$\int_a^b (f \cdot G)(x)dx = G(a) \int_a^\xi f(x)dx. \tag{29}$$

但

$$\int_a^b (f \cdot G)(x)dx = g(b) \int_a^b f(x)dx - \int_a^b (f \cdot g)(x)dx,$$

$$G(a) \int_a^\xi f(x)dx = g(b) \int_a^\xi f(x)dx - g(a) \int_a^\xi f(x)dx.$$

注意到这些关系式以及积分的可加性, 从 (29) 即可得到待证的等式 (28).

如果 g 是不增函数, 那么令 $G(x) = g(x) - g(b)$, 我们就得到 $G(x)$ 是 $[a, b]$ 上的非负, 不增, 可积的函数. 这样, 我们又得到公式 (29), 跟着也就得出公式 (28). ►

---

①波内(Bonnet)(1819—1892)—— 法国数学家和天文学家 波内最主要的数学著作是关于微分几何方面的.

练 习

1. 试证: 如果 $f \in \mathcal{R}[a,b]$ 且在 $[a,b]$ 上有 $f(x) \geqslant 0$, 则

a) 当在 $f(x)$ 的某一连续点 $x_0 \in [a,b]$ 有 $f(x_0) > 0$ 时, 必成立严格不等式

$$\int_a^b f(x)dx > 0.$$

b) 从条件 $\int_a^b f(x)dx = 0$ 可推出在 $[a,b]$ 上几乎处处成立 $f(x) = 0$.

2. 试证: 如果 $f \in \mathcal{R}[a,b], m = \inf\limits_{x\in]a,b[} f(x), M = \sup\limits_{x\in]a,b[} f(x)$, 则

a) $\int_a^b f(x)dx = \mu(b-a)$, 其中 $\mu \in [m, M]$(见上节问题 5a)).

b) 当 f 在 $[a,b]$ 上连续时, 存在点 $\xi \in]a,b[$ 使

$$\int_a^b f(x)dx = f(\xi)(b-a).$$

3. 试证: 如果 $f \in C[a,b], f(x) \geqslant 0$ 在 $[a,b]$ 上成立, 以及 $M = \max\limits_{x\in[a,b]} f(x)$, 则

$$\lim_{n\to+\infty} \left[\int_a^b f^n(x)dx\right]^{\frac{1}{n}} = M.$$

4. a) 试证: 如果 $f \in \mathcal{R}[a,b]$, 则 $|f|^p \in \mathcal{R}[a,b] (p \geqslant 0)$.

b) 试根据和式的赫尔德不等式证明积分的赫尔德不等式 ①:

$$\left|\int_a^b (f\cdot g)(x)dx\right| \leqslant \left[\int_a^b |f|^p(x)dx\right]^{\frac{1}{p}} \left[\int_a^b |g|^q(x)dx\right]^{\frac{1}{q}},$$

其中 $f, g \in \mathcal{R}[a,b]$ 且 $p \geqslant 1, q \geqslant 1, \dfrac{1}{p} + \dfrac{1}{q} = 1$.

c) 试根据和式的闵可夫斯基不等式证明积分的闵可夫斯基不等式:

$$\left[\int_a^b |f+g|^p(x)dx\right]^{\frac{1}{p}} \leqslant \left[\int_a^b |f|^p(x)dx\right]^{\frac{1}{p}} + \left[\int_a^b |g|^p(x)dx\right]^{\frac{1}{p}},$$

其中 $f, g \in \mathcal{R}[a,b]$ 且 $p \geqslant 1$. 试证: 当 $0 < p < 1$ 时, 这个不等式的不等号调转方向.

d) 试验证. 如果 f 是 $\mathbb{R}$ 上的连续凸函数, 而 φ 是 $\mathbb{R}$ 上的任意连续函数, 则当 $c \neq 0$ 时成立詹生不等式

$$f\left(\frac{1}{c}\int_0^c \varphi(t)dt\right) \leqslant \frac{1}{c}\int_0^c f(\varphi(t))dt.$$

---

① 当 $p = q = 2$ 时, 赫尔德代数不等式首先是柯西于 1821 年得到的, 并以他的名字命名. 当 $p = q = 2$ 时, 赫尔德积分不等式首先是俄国数学家布尼雅可夫斯基 (1804—1889) 于 1859 年发现的. 这个重要不等式 ($p = q = 2$ 的情形) 叫做布尼雅可夫斯基不等式或柯西-布尼雅可夫斯基不等式. 有时还看到不准确的命名——"施瓦茨不等式". 在德国数学家施瓦茨(Schwartz) (1843—1921) 1884 年的著作中出现过这个不等式.

§3. 积分和导数

1. 积分和原函数

设 f 是闭区间 $[a,b]$ 上的黎曼可积函数. 考虑这个区间上的函数

$$F(x)=\int_a^x f(t)dt, \tag{1}$$

常把它叫做变上限的积分.

因为 $f\in\mathcal{R}[a,b]$, 所以 $f|_{[a,x]}\in\mathcal{R}[a,x]$, 其中 $[a,x]\subset[a,b]$; 因此, 函数 $x\mapsto F(x)$ 对 $x\in[a,b]$ 有定义.

如果 $|f(x)|\leqslant C<+\infty$ 在 $[a,b]$ 上成立 (f 作为可积函数在 $[a,b]$ 上有界), 则从积分的可加性和最简单的估计得到: 如果 x、$x+h\in[a,b]$, 则

$$|F(x+h)-F(x)|\leqslant C|h|. \tag{2}$$

顺便指出, 关于这一点在证明上节的引理 3 时我们已经说过.

特别地, 从不等式 (2) 得到 F 在 $[a,b]$ 上连续. 于是 $F\in C[a,b]$.

现在, 我们对 F 进行更详细的研究.

下面的引理对以后的整个理论具有根本的重要性.

引理 1　*如果 $f\in\mathcal{R}[a,b]$, 而 f 在某点 $x\in[a,b]$ 连续, 则公式 (1) 定义的函数 F 在这个点可微, 而且成立等式*

$$F'(x)=f(x).$$

◀ 设 $x,x+h\in[a,b]$. 我们来估计差 $F(x+h)-F(x)$. 从 f 在 x 的连续性推出 $f(t)=f(x)+\Delta(t)$, 其中当 $t\to x$ 时有 $\Delta(t)\to 0, t\in[a,b]$. 如果 x 是固定点, 则函数 $\Delta(t)=f(t)-f(x)$ 作为可积函数 $t\mapsto f(t)$ 和常数 $f(x)$ 之差是区间 $[a,b]$ 上的可积函数, 以 $M(h)$ 表示量 $\sup\limits_{t\in I(h)}|\Delta(t)|$, 其中 $I(h)$ 是以 x 和 $x+h\in[a,b]$ 为端点的区间. 根据假设条件, 当 $h\to 0$ 时, 有 $M(h)\to 0$.

记

$$\begin{aligned}F(x+h)-F(x)&=\int_a^{x+h}f(t)dt-\int_a^x f(t)dt=\int_x^{x+h}f(t)dt\\&=\int_x^{x+h}[f(x)+\Delta(t)]dt=\int_x^{x+h}f(x)dt+\int_x^{x+h}\Delta(t)dt\\&=f(x)h+\alpha(h)h,\end{aligned}$$

其中

$$\int_x^{x+h}\Delta(t)dt=\alpha(h)h.$$

因为

$$\left|\int_x^{x+h}\Delta(t)dt\right| \leqslant \left|\int_x^{x+h}|\Delta(t)|dt\right| \leqslant \left|\int_x^{x+h}M(h)dt\right| = M(h)|h|,$$

所以 $|\alpha(h)| \leqslant M(h)$, 从而, 当 $h \to 0$ 时 (但 $x+h \in [a,b]$), 有 $\alpha(h) \to 0$.

这样一来, 我们证明了: 如果 f 在点 $x \in [a,b]$ 连续, 则当位移 h 使 $x+h \in [a,b]$ 时, 成立等式

$$F(x+h) - F(x) = f(x)h + \alpha(h)h, \tag{3}$$

其中 $\alpha(h) \to 0$, 当 $h \to 0$ 时.

然而, 这个结果正说明函数 F 在点 $x \in [a,b]$ 处可微且 $F'(x) = f(x)$. ►

引理 1 的最重要的直接推论是下面的定理.

定理 1 *闭区间 $[a,b]$ 上的每个连续函数 $f:[a,b] \to \mathbb{R}$ 在该区间上都有一个原函数, 而且区间 $[a,b]$ 上的函数 f 的任一原函数都有*

$$\mathcal{F}(x) = \int_a^x f(t)dt + c \tag{4}$$

的形式, 其中 c 是一个常数.

◄ $(f \in C[a,b]) \Rightarrow (f \in \mathcal{R}[a,b])$, 因此, 根据引理 1, 函数 (1) 是 f 在 $[a,b]$ 上的原函数. 但是, 同一个函数在区间上的两个原函数 $\mathcal{F}(x)$ 和 $F(x)$ 只能相差一个常数, 因此, $\mathcal{F}(x) = F(x) + c$. ►

为了今后应用的方便, 我们把原函数概念作稍许推广, 采用

定义 1 区间上的连续函数 $x \mapsto \mathcal{F}(x)$ 叫做定义在该区间上的函数 $x \mapsto f(x)$ 的*原函数 (广义原函数)*, 如果最多除去有限多个点以外, 在该区间上成立关系 $\mathcal{F}'(x) = f(x)$.

根据这个定义可以证明:

定理 1′ *在闭区间 $[a,b]$ 上定义的有界且仅有有限多个间断点的函数 $f:[a,b] \to \mathbb{R}$, 在该区间上有 (广义) 原函数, 而且 f 在 $[a,b]$ 上的任一原函数都具有 (4) 的形式.*

◄ 因为 f 只有有限个间断点且有界, 所以 $f \in \mathcal{R}[a,b]$, 从而, 根据引理 1, 函数 (1) 是 f 在 $[a,b]$ 上的广义原函数, 而且, 如已经指出的那样, 根据 (2), 函数 (1) 在 $[a,b]$ 上连续. 如果 $\mathcal{F}(x)$ 是函数 f 在 $[a,b]$ 上的另一个原函数, 则 $\mathcal{F}(x) - F(x)$ 是连续的, 而且在由 f 的间断点分割区间 $[a,b]$ 而成的每个开区间内是常数. 由此, 从 $\mathcal{F}(x) - F(x)$ 在 $[a,b]$ 上的连续性得出, 在 $[a,b]$ 上 $\mathcal{F}(x) - F(x) \equiv \text{const}$. ►

2. 牛顿–莱布尼茨公式

定理 2　如果 $f:[a,b]\to\mathbb{R}$ 是有界且仅有有限个间断点的函数, 则 $f\in\mathcal{R}[a,b]$ 且

$$\boxed{\int_a^b f(x)dx=\mathcal{F}(b)-\mathcal{F}(a),}\tag{5}$$

其中 $\mathcal{F}:[a,b]\to\mathbb{R}$ 是 f 在闭区间 $[a,b]$ 上的任一原函数.

◀ 在闭区间上定义的仅有有限个间断点的有界函数的可积性是已经证明过的(见 §1 命题 2 的推论 2). 函数 f 在 $[a,b]$ 上的原函数的存在性由定理 1′ 保证, 根据这个定理, $\mathcal{F}(x)$ 有 (4) 的形式. 在 (4) 中令 $x=a$ 得 $F(a)=c$, 由此

$$\mathcal{F}(x)=\int_a^x f(t)dt+\mathcal{F}(a).$$

特别地,

$$\int_a^b f(t)dt=\mathcal{F}(b)-\mathcal{F}(a),$$

这与要证的 (5) 只是表示积分变量的字母不同而已. ▶

关系 (5) 对整个分析学具有基本的意义, 它叫做**牛顿–莱布尼茨公式**.

函数的差值 $\mathcal{F}(b)-\mathcal{F}(a)$ 常用符号 $\mathcal{F}(x)|_a^b$ 表示, 利用这个记号可将牛顿–莱布尼茨公式写成如下形式:

$$\boxed{\int_a^b f(x)dx=\mathcal{F}(x)\Big|_a^b.}$$

由于当 a 和 b 调换时, 这个公式两端同时改变符号, 所以该公式对于任何关系的 a 和 b, 也就是, 无论 $a\leqslant b$ 还是 $a\geqslant b$, 都是正确的.

在分析练习中, 牛顿–莱布尼茨公式多用来计算它左边的积分, 这容易造成对这个公式应用的误解. 事实上, 具体的积分很少是通过原函数求出的, 而常常是借助于常用的积分近似计算法用电子计算机直接进行数值计算求出的. 牛顿–莱布尼茨公式在数学分析理论本身中占有关键地位, 它把积分学和微分学联系起来. 特别是, 它在分析中还进一步发展成一般的**斯托克斯**[①]公式.

这一节剩下的部分可以看作是牛顿–莱布尼茨公式在分析本身中应用的例子.

3. 定积分的分部积分法和泰勒公式

命题 1　如果函数 $u(x)$ 和 $v(x)$ 在以 a 和 b 为端点的闭区间上连续可微, 则成立关系

$$\int_a^b(u\cdot v')(x)dx=(u\cdot v)(x)\Big|_a^b-\int_a^b(v\cdot u')(x)dx.\tag{6}$$

①斯托克斯(G. G. Stokes) (1819—1903)　英国物理学家和数学家.

习惯上常把这个公式写成以下简略形式:

$$\int_a^b u dv = u \cdot v\Big|_a^b - \int_a^b v du,$$

称它为定积分的分部积分公式.

◀ 根据函数乘积的微分规则, 我们有

$$(u \cdot v)'(x) = (u' \cdot v)(x) + (u \cdot v')(x).$$

根据假设条件, 这个等式中的函数都是连续的, 从而在以 a 和 b 为端点的区间上可积. 利用积分的线性性和牛顿–莱布尼茨公式, 我们得到

$$(u \cdot v)(x)\Big|_a^b = \int_a^b (u' \cdot v)(x)dx + \int_a^b (u \cdot v')(x)dx. \ ▶$$

作为推论, 我们将得到带积分余项的泰勒公式.

设以 a 和 x 为端点的闭区间上的函数 $t \mapsto f(t)$ 有 n 阶连续导数. 利用牛顿–莱布尼茨公式和公式 (6), 进行下面一串演算, 其中所有的微分和代入都是关于变量 t 做的:

$$\begin{aligned}
f(x) - f(a) &= \int_a^x f'(t)dt = -\int_a^x f'(t)(x-t)'dt \\
&= -f'(t)(x-t)\Big|_a^x + \int_a^x f''(t)(x-t)dt \\
&= f'(a)(x-a) - \frac{1}{2}\int_a^x f''(t)((x-t)^2)'dt \\
&= f'(a)(x-a) - \frac{1}{2}f''(t)(x-t)^2\Big|_a^x + \frac{1}{2}\int_a^x f'''(t)(x-t)^2dt \\
&= f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 - \frac{1}{2\cdot 3}\int_a^x f'''(t)((x-t)^3)'dt \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
&= f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \cdots + \\
&\quad \frac{1}{2\cdot 3\cdot \cdots \cdot (n-1)}f^{(n-1)}(a)(x-a)^{n-1} + r_n(a;x),
\end{aligned}$$

其中

$$\boxed{r_n(a,x) = \frac{1}{(n-1)!}\int_a^x f^{(n)}(t)(x-t)^{n-1}dt.} \tag{7}$$

这样, 我们就证明了

命题 2 如果函数 $t \mapsto f(t)$ 在以 a 和 x 为端点的闭区间上有直到 n 阶连续导数, 则成立泰勒公式

$$f(x) = f(a) + \frac{1}{1!}f'(a)(x-a) + \cdots + \frac{1}{(n-1)!}f^{(n-1)}(a)(x-a)^{n-1} + r_n(a;x)$$

其中 $r_n(a;x)$ 是形如 (7) 的积分形式泰勒余项.

我们指出, 在以 a 和 x 为端点的闭区间上函数 $(x-t)^{n-1}$ 不改变符号, 而函数 $t \mapsto f^{(n)}(t)$ 在这个区间上连续. 因此, 根据第一中值定理在这区间上存在点 ξ 使

$$\begin{aligned} r_n(a;x) &= \frac{1}{(n-1)!}\int_a^x f^{(n)}(t)(x-t)^{n-1}dt \\ &= \frac{1}{(n-1)!}f^{(n)}(\xi)\int_a^x (x-t)^{n-1}dt \\ &= \frac{1}{(n-1)!}f^{(n)}(\xi)\left(-\frac{1}{n}(x-t)^n\right)\Big|_a^x \\ &= \frac{1}{n!}f^{(n)}(\xi)(x-a)^n. \end{aligned}$$

这样, 我们又导出了熟知的拉格朗日形式的泰勒公式余项 (根据上一节的问题 2b), 可以认为 ξ 在以 a 和 x 为端点的开区间内.)

如果从 (7) 式的积分号下拿出 $f^{(n)}(\xi)(x-\xi)^{n-k}(k \in [1,n])$, 上边的推理可照样进行. 同时, 与 $k=1$ 和 $k=n$ 相对应, 得到余项的柯西公式和拉格朗日公式.

4. 定积分中的变量替换 定积分中的变量替换公式是积分学中的基本公式之一. 这个公式在积分理论中, 就像复合函数的微分公式在微分学中一样地重要, 并在一定条件下借助于牛顿-莱布尼茨公式与其相联系.

命题 3 如果 $\varphi: [\alpha,\beta] \to [a,b]$ 是从闭区间 $\alpha \leqslant t \leqslant \beta$ 到闭区间 $a \leqslant x \leqslant b$ 的连续可微映射, 且 $\varphi(\alpha) = a$ 和 $\varphi(\beta) = b$, 则对于在 $[a,b]$ 上连续的任何函数 $f(x)$, 函数 $f(\varphi(t))\varphi'(t)$ 在闭区间 $[\alpha,\beta]$ 上连续且成立等式

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)dt. \tag{8}$$

◀ 设 $\mathcal{F}(x)$ 是函数 $f(x)$ 在 $[a,b]$ 上的原函数, 则根据复合函数微分法定理知道, $\mathcal{F}(\varphi(t))$ 是函数 $f(\varphi(t))\varphi'(t)$ 的原函数, 因为函数 $f(\varphi(t))\varphi'(t)$ 是区间 $[\alpha,\beta]$ 上的连续函数复合和乘积. 根据牛顿-莱布尼茨公式, 有

$$\int_a^b f(x)dx = \mathcal{F}(b) - \mathcal{F}(a)$$

和

$$\int_\alpha^\beta f(\varphi(t))\varphi'(t)dt = \mathcal{F}(\varphi(\beta)) - \mathcal{F}(\varphi(\alpha)).$$

但根据条件 $\varphi(\alpha)=a$ 和 $\varphi(\beta)=b$, 可得等式 (8) 成立. ►

从公式 (8) 看出, 在积分号下不用函数符号而用微分表达式 $f(x)dx$ 是有道理的, 它使得在经过变换 $x=\varphi(t)$ 后能自动地得到关于新变量的积分号下的表达式.

为了不致使事情因烦琐的证明而复杂化, 在命题 3 中我们故意缩小了公式 (8) 的使用范围, 从而可由牛顿-莱布尼茨公式得到它. 现在我们转入变量替换基本定理的讨论, 它的条件与命题 3 的条件有些不同. 这个定理的证明直接从积分的定义出发, 即把积分作为积分和的极限.

定理 3 *设 $\varphi:[\alpha,\beta]\to[a,b]$ 是从闭区间 $\alpha\leqslant t\leqslant\beta$ 到闭区间 $a\leqslant x\leqslant b$ 的连续可微且严格单调映射, 而且 $\varphi(\alpha)=a,\varphi(\beta)=b$ 或 $\varphi(\alpha)=b,\varphi(\beta)=a$ 那么, 对于在区间 $[a,b]$ 上可积的任意函数 $f(x)$, 函数 $f(\varphi(t))\varphi'(t)$ 在区间 $[\alpha,\beta]$ 上可积且成立等式:*

$$\boxed{\int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)dx=\int_\alpha^\beta f(\varphi(t))\varphi'(t)dt.} \tag{9}$$

◄ 由于 φ 是从区间 $[\alpha,\beta]$ 到区间 $[a,b]$ 上的严格单调函数, 所以区间 $[\alpha,\beta]$ 的任意一个分划 $P_t(\alpha=t_0<\cdots<t_n=\beta)$ 借助于分划 P_t 的点的象 $x_i=\varphi(t_i)(i=0,1,\cdots,n)$ 产生区间 $[a,b]$ 的一个相应的分划 P_x, 我们把它记做 $\varphi(P_t)$. 同时, 当 $\varphi(\alpha)=a$ 时有 $x_0=a$, 而当 $\varphi(\alpha)=b$ 时有 $x_0=b$. 由 φ 在 $[\alpha,\beta]$ 上的一致连续性得到, 如果 $\lambda(P_t)\to0$, 则 $\lambda(P_x)=\lambda(\varphi(P_t))$ 也趋于零.

利用拉格朗日定理, 把积分和 $\sigma(f;P_x,\xi)$ 变成以下形式:

$$\begin{aligned}\sum_{i=1}^n f(\xi_i)\Delta x_i&=\sum_{i=1}^n f(\xi_i)(x_i-x_{i-1})\\&=\sum_{i=1}^n f(\varphi(\tau_i))\varphi'(\tilde{\tau}_i)(t_i-t_{i-1})\\&=\sum_{i=1}^n f(\varphi(\tau_i))\varphi'(\tilde{\tau}_i)\Delta t_i.\end{aligned}$$

这里 $x_i=\varphi(t_i),\xi_i=\varphi(\tau_i),\xi_i$ 属于以 x_{i-1},x_i 为端点的区间, τ_i 属于以 t_{i-1},t_i 为端点的区间 $(i=1,\cdots,n)$.

其次

$$\begin{aligned}\sum_{i=1}^n f(\varphi(\tau_i))\varphi'(\tilde{\tau}_i)\Delta t_i=&\sum_{i=1}^n f(\varphi(\tau_i))\varphi'(\tau_i)\Delta t_i\\&+\sum_{i=1}^n f(\varphi(\tau_i))(\varphi'(\tilde{\tau}_i)-\varphi'(\tau_i))\Delta t_i.\end{aligned}$$

我们来估计最后这个和. 由于 $f \in \mathcal{R}[a,b]$, 函数 f 在 $[a,b]$ 上有界. 设 $|f(x)| \leqslant C$ 在 $[a,b]$ 上成立. 那么,

$$\left|\sum_{i=1}^{n} f(\varphi(\tau_i))(\varphi'(\tilde{\tau}_i) - \varphi'(\tau_i))\Delta t_i\right| \leqslant C \cdot \sum_{i=1}^{n} \omega(\varphi'; \Delta_i)\Delta t_i,$$

这里 Δ_i 表示以 t_{i-1}, t_i 为端点的区间.

上式右端的和式当 $\lambda(P_t) \to 0$ 时趋于零, 因为 φ' 是区间 $[\alpha,\beta]$ 上的连续函数.

这样一来, 我们证明了:

$$\sum_{i=1}^{n} f(\xi_i)\Delta x_i = \sum_{i=1}^{n} f(\varphi(\tau_i))\varphi'(\tau_i)\Delta t_i + \alpha,$$

其中 α 当 $\lambda(P_t) \to 0$ 时趋于零. 已经看到, 如果 $\lambda(P_t) \to 0$, 就有 $\lambda(P_x) \to 0$, 但 $f \in \mathcal{R}[a,b]$, 因此, 当 $\lambda(P_x) \to 0$ 时, 上面等式的左端趋于积分 $\int_{\varphi(\alpha)}^{\varphi(\beta)} f(x)dx$. 因此, 当 $\lambda(P_t) \to 0$ 时, 这个等式右端应有相同的极限.

然而, 和 $\sum\limits_{i=1}^{n} f(\varphi(\tau_i))\varphi'(\tau_i)\Delta t_i$ 完全可以看作是对应于带标志点 $\tau = (\tau_1, \cdots, \tau_n)$ 的分划 P_t 的一个积分和, 这是因为由 φ 的严格单调性, 任何点组 τ 都能从分划 $P_x = \varphi(P_t)$ 的区间中与之相应的标志点组 $\xi = (\xi_1, \cdots, \xi_n)$ 得到.

这样一来, 根据定义, 这个和式的极限是函数 $f(\varphi(t))\varphi'(t)$ 在区间 $[\alpha,\beta]$ 上的积分, 这就同时证明了函数 $f(\varphi(t))\varphi'(t)$ 在区间 $[\alpha,\beta]$ 上的可积性和公式 (9). ►

5. 一些例子 现在考虑几个例子, 其中将要用到刚讲过的这两节中所得到的公式和证明过的关于积分性质的定理.

例 1

$$\begin{aligned}\int_{-1}^{1}\sqrt{1-x^2}dx &= \int_{-\pi/2}^{\pi/2}\sqrt{1-\sin^2 t}\cos t dt \\ &= \int_{-\pi/2}^{\pi/2}\cos^2 t dt = \frac{1}{2}\int_{-\pi/2}^{\pi/2}(1+\cos 2t)dt \\ &= \frac{1}{2}\left(t + \frac{1}{2}\sin 2t\right)\Big|_{-\pi/2}^{\pi/2} = \frac{\pi}{2}.\end{aligned}$$

为了计算这个积分, 我们做了变量替换 $x = \sin t$, 然后求出经过这个变换后的被积函数的原函数, 再应用牛顿–莱布尼茨公式.

当然, 也可以用另外的方法去做: 首先求出函数 $\sqrt{1-x^2}$ 的相当复杂的原函数 $\frac{1}{2}x\sqrt{1-x^2} + \frac{1}{2}\arcsin x$, 然后再用牛顿–莱布尼茨公式. 这个例子说明, 在计算定积分时, 有时能幸运地避开求复杂的被积函数的原函数.

例 2 证明: 当 $m,n\in\mathbf{N}$ 时, 有

a) $\displaystyle\int_{-\pi}^{\pi}\sin mx\cos nx dx=0,$

b) $\displaystyle\int_{\pi}^{\pi}\sin^2 mx dx=\pi,$

c) $\displaystyle\int_{-\pi}^{\pi}\cos^2 nx dx=\pi.$

◀ a) $\displaystyle\int_{-\pi}^{\pi}\sin mx\cos nx dx$

$$=\frac{1}{2}\int_{-\pi}^{\pi}(\sin(n+m)x-\sin(n-m)x)dx$$

$$=\frac{1}{2}\left(-\frac{1}{n+m}\cos(m+n)x+\frac{1}{n-m}\cos(n-m)x\right)\Big|_{-\pi}^{\pi}=0,$$

其中 $n-m\neq 0$. 而当 $n-m=0$ 时可以单独计算, 显然这时也得到同样结果.

b) $\displaystyle\int_{-\pi}^{\pi}\sin^2 mx dx=\frac{1}{2}\int_{-\pi}^{\pi}(1-\cos 2mx)dx=\frac{1}{2}\left(x-\frac{1}{2m}\sin 2mx\right)\Big|_{-\pi}^{\pi}=\pi.$

c) $\displaystyle\int_{-\pi}^{\pi}\cos^2 nx dx=\frac{1}{2}\int_{-\pi}^{\pi}(1+\cos 2nx)dx=\frac{1}{2}\left(x+\frac{1}{2n}\sin 2nx\right)\Big|_{-\pi}^{\pi}=\pi.$ ▶

例 3 设 $f\in\mathcal{R}[-a,a]$. 证明:

$$\int_{-a}^{a}f(x)dx=\begin{cases}2\displaystyle\int_0^a f(x)dx, & \text{当}f\text{是偶函数},\\ 0, & \text{当}f\text{是奇函数}.\end{cases}$$

如果 $f(-x)=f(x)$, 则

$$\begin{aligned}\int_{-a}^{a}f(x)dx&=\int_{-a}^{0}f(x)dx+\int_0^a f(x)dx\\&=\int_a^0 f(-t)\cdot(-1)dt+\int_0^a f(x)dx\\&=\int_0^a f(-t)dt+\int_0^a f(x)dx\\&=\int_0^a (f(x)+f(-x))dx=2\int_0^a f(x)dx.\end{aligned}$$

如果 $f(-x)=-f(x)$, 仍然通过上面那些计算可以看出

$$\int_{-a}^{a}f(x)dx=\int_0^a(f(x)+f(-x))dx=\int_0^a 0\cdot dx=0.$$

例 4 设 f 是在整个数轴 $\mathbb{R}$ 上定义的以 T 为周期的周期函数, 即 $f(x+T)=f(x)$ 对于 $x\in\mathbb{R}$.

如果 f 在每个有限区间上可积, 则对于任何 $a \in \mathbb{R}$ 成立不等式

$$\int_a^{a+T} f(x)dx = \int_a^T f(x)dx,$$

也就是说, 周期函数在长度等于它的周期的区间上的积分不依赖于积分区间在数轴上的位置:

$$\begin{aligned}\int_a^{a+T} f(x)dx &= \int_a^0 f(x)dx + \int_0^T f(x)dx + \int_T^{a+T} f(x)dx \\ &= \int_0^T f(x)dx + \int_a^0 f(x)dx + \int_0^a f(t+T)\cdot 1 \cdot dt \\ &= \int_0^T f(x)dx + \int_a^0 f(x)dx + \int_0^a f(t)dt = \int_0^T f(x)dx.\end{aligned}$$

在这里我们做了替换 $x = t + T$, 并应用了函数 $f(x)$ 的周期性.

例 5 假设需要计算积分 $\displaystyle\int_0^1 \sin x^2 dx$ 精确到 10^{-2} 的值.

我们知道, 函数 $\sin x^2$ 的原函数 $\displaystyle\int \sin x^2 dx$ (菲涅耳积分) 不能表示成初等函数, 因此, 像通常那样地利用牛顿 莱布尼茨公式, 在这里是不行的.

我们用另外的做法. 在微分学中研究泰勒公式时, 作为例子我们曾指出 (见第五章 §3 的例 11): 在区间 $[-1,1]$ 上成立精确到 10^{-3} 的等式

$$\sin x \approx x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 = P(x).$$

然而, 如果在区间 $[-1,1]$ 上有 $|\sin x - P(x)| < 10^{-3}$, 则成立 $|\sin x^2 - P(x^2)| < 10^{-3}$ 当 $0 \leqslant x \leqslant 1$. 因此,

$$\begin{aligned}\left|\int_0^1 \sin x^2 dx - \int_0^1 P(x^2)dx\right| &\leqslant \int_0^1 |\sin x^2 - P(x^2)|dx \\ &< \int_0^1 10^{-3}dx = 10^{-3}.\end{aligned}$$

这样一来, 为了以所需要的精确度计算出 $\displaystyle\int_0^1 \sin x^2 dx$ 的值, 只要计算 $\displaystyle\int_0^1 P(x^2)dx$ 就够了. 但是,

$$\begin{aligned}\int_0^1 P(x^2)dx &= \int_0^1 \left(x^2 - \frac{1}{3!}x^6 + \frac{1}{5!}x^{10}\right)dx \\ &= \left(\frac{1}{3}x^3 - \frac{1}{3!7}x^7 + \frac{1}{5!11}x^{11}\right)\Big|_0^1 \\ &= \frac{1}{3} - \frac{1}{3!7} + \frac{1}{5!11} = 0.310 \pm 10^{-3}.\end{aligned}$$

因此,

$$\int_0^1 \sin x^2 dx = 0.310 \pm 2 \cdot 10^{-3} = 0.31 \pm 10^{-2}.$$

例 6 值 $\mu = \frac{1}{b-a}\int_a^b f(x)dx$ 叫做函数 $f(x)$ 在区间 $[a,b]$ 上的积分平均值.

设 f 是在 $\mathbb{R}$ 上定义且在任何区间上都可积的函数.

用 f 构造新函数

$$F_\delta(x) = \frac{1}{2\delta}\int_{x-\delta}^{x+\delta} f(t)dt,$$

它在点 x 处的值是 f 在点 x 的 δ-邻域中的积分平均值.

我们来证明, 函数 $F_\delta(x)$(它叫 f 的平均函数) 比 f 更正则些. 确切地说, 如果 f 在任何区间 $[a,b]$ 上都可积, 则 F_δ 在 $\mathbb{R}$ 上连续; 而如果 $f \in C(\mathbb{R})$, 则 $F_\delta \in C^1(\mathbb{R})$.

首先检验函数 $F_\delta(x)$ 的连续性:

$$|F_\delta(x+h) - F_\delta(x)| = \frac{1}{2\delta}\left|\int_{x+\delta}^{x+\delta+h} f(t)dt + \int_{x-\delta+h}^{x-\delta} f(t)dt\right|$$
$$\leqslant \frac{1}{2\delta}(C|h| + C|h|) = \frac{C}{\delta}|h|,$$

这里假设 $|f(t)| \leqslant C$, 为此只要 t 在 x 的 2δ-邻域中变动且 $|h| < \delta$ 即可. 显然, 从这个估计立刻推出函数 $F_\delta(x)$ 的连续性.

如果 $f \in C(\mathbb{R})$, 根据复合函数微分法, 有

$$\frac{d}{dx}\int_a^{\varphi(x)} f(t)dt = \frac{d}{d\varphi}\int_a^{\varphi} f(t)dt \cdot \frac{d\varphi}{dx} = f(\varphi(x))\varphi'(x).$$

因此, 从

$$F_\delta(x) = \frac{1}{2\delta}\int_a^{x+\delta} f(t)dt - \frac{1}{2\delta}\int_a^{x-\delta} f(t)dt$$

得到

$$F'_\delta(x) = \frac{f(x+\delta) - f(x-\delta)}{2\delta}.$$

经过积分变量的替换 $t = x + u$ 之后, 函数 $F_\delta(x)$ 可以写成以下形式:

$$F_\delta(x) = \frac{1}{2\delta}\int_{-\delta}^{\delta} f(x+u)du.$$

如果 $f \in C(\mathbb{R})$, 利用第一中值定理, 就得到

$$F_\delta(x) = \frac{1}{2\delta} f(x+\tau) \cdot 2\delta = f(x+\tau),$$

其中 $|\tau| \leqslant \delta$. 由此推出很自然的结果:

$$\lim_{\delta \to +0} F_\delta(x) = f(x).$$

练　习

1. 利用积分求

a) $\lim\limits_{n\to+\infty}\left[\dfrac{n}{(n+1)^2}+\cdots+\dfrac{n}{(2n)^2}\right]$.

b) $\lim\limits_{n\to\infty}\dfrac{1^\alpha+2^\alpha+\cdots+n^\alpha}{n^{\alpha+1}}, \alpha\geqslant 0$.

2. a) 试证: 开区间上的任何连续函数在该区间上都有原函数.

b) 试证: 如果 $f\in C^{(1)}[a,b]$, 则 f 可以表示成区间 $[a,b]$ 上的两个不减函数之差 (见 §1 的问题 4).

3. 试证: 在函数 g 光滑假定下, 利用分部积分即可把积分第二中值定理 (§2 的定理 6) 直接归结到第一中值定理.

4. 试证: 如果 $f\in C(\mathbb{R})$, 则对于任意确定的区间 $[a,b]$, 根据给定的 $\varepsilon>0$, 可以取 $\delta>0$, 使在区间 $[a,b]$ 上成立不等式 $|F_\delta(x)-f(x)|<\varepsilon$, 其中 F_δ 是例 6 中的平均函数.

5. 试证: 当 $x\to+\infty$ 时, 有

$$\int_1^{x^2}\frac{e^t}{t}dt\sim\frac{2}{x}e^{x^2}.$$

6. a) 验证: 当 $x\to\infty$ 时, 函数 $f(x)=\displaystyle\int_x^{x+1}\sin t^2dt$ 有下列表示

$$f(x)=\frac{\cos x^2}{x}-\frac{\cos(x+1)^2}{x+1}+O\left(\frac{1}{x^2}\right).$$

b) 求 $\underline{\lim\limits_{x\to\infty}}\,xf(x)$ 和 $\overline{\lim\limits_{x\to\infty}}\,xf(x)$.

7. 试证: 如果 $f:\mathbb{R}\to\mathbb{R}$ 是在每个闭区间 $[a,b]\subset\mathbb{R}$ 上可积的周期函数, 则函数

$$F(x)=\int_a^x f(t)dt$$

可以表示成线性函数与周期函数之和的形式.

8. a) 验证: 当 $x>1$ 和 $n\in\mathbb{N}$ 时, 函数

$$P_n(x)=\frac{1}{\pi}\int_0^\pi(x+\sqrt{x^2-1}\cos\varphi)^n d\varphi$$

是 n 次多项式 (n 次勒让德多项式).

b) 试证:

$$P_n(x)=\frac{1}{\pi}\int_0^\pi\frac{d\psi}{(x-\sqrt{x^2-1}\cos\psi)^n}.$$

9. 设 f 是定义在区间 $[a,b]\subset\mathbb{R}$ 上的实值函数, 而 $\xi_1,\cdots,\xi_m$ 是这个区间的不同点. $m-1$ 阶拉格朗日插值多项式

$$L_{m-1}(x)=\sum_{j=1}^m f(\xi_j)\prod_{i\neq j}\frac{x-\xi_j}{\xi_j-\xi_i}$$

的值与函数 f 的值在点 $\xi_1, \cdots, \xi_m$(插值节点) 重合, 而且, 如果 $f \in C^{(m)}[a,b]$, 则

$$f(x) - L_{m-1}(x) = \frac{1}{m!} f^{(m)}(\zeta(x))\omega_m(x),$$

其中 $\omega_m(x) = \prod_{i=1}^{m}(x - \xi_i)$, 而 $\zeta(x) \in]a,b[$(见第五章 §3 的问题 11).

设 $\xi_i = \frac{b+a}{2} + \frac{b-a}{2}\theta_i$, 则 $\theta_i \in [-1,1] (i = 1, \cdots, m)$.

a) 试证:

$$\int_a^b L_{m-1}(x)dx = \frac{b-a}{2}\sum_{i=1}^{m} c_i f(\xi_i),$$

其中

$$c_i = \int_{-1}^{1}\left(\prod_{j \neq i}\frac{t-\theta_j}{\theta_j - \theta_i}\right)dt.$$

特别地,

$\alpha_1)$ $\int_a^b L_0(x)dx = (b-a)f\left(\frac{a+b}{2}\right)$, 当 $m = 1, \theta_1 = 0$;

$\alpha_2)$ $\int_a^b L_1(x)dx = \frac{b-a}{2}[f(a) + f(b)]$, 当 $m = 2, \theta_1 = -1, \theta_2 = 1$;

$\alpha_3)$ $\int_a^b L_2(x)dx = \frac{b-a}{6}\left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right]$, 当 $m = 3, \theta_1 = -1, \theta_2 = 0, \theta_3 = 1$.

b) 设 $f \in C^{(m)}[a,b]$, 令 $M_m = \max_{x \in [a,b]}|f^{(m)}(x)|$. 估计公式

$$\int_a^b f(x)dx = \int_a^b L_{m-1}(x)dx + R_m \qquad (*)$$

中 R_m 的绝对误差, 并证明 $|R_m| \leqslant \frac{M_m}{m!}\int_a^b |\omega_m(x)|dx$.

c) 在情形 $\alpha_1), \alpha_2), \alpha_3)$ 中公式 (*) 依次称做矩形公式, 梯形公式和抛物形公式. 在最后一种情形中, 这个公式也叫辛普森公式 [①]. 试证: 在情形 $\alpha_1), \alpha_2), \alpha_3)$ 中成立公式:

$$R_1 = \frac{f'(\xi_1)}{4}(b-a)^2, R_2 = -\frac{f''(\xi_2)}{12}(b-a)^3, R_3 = -\frac{f^{(4)}(\xi_3)}{2880}(b-a)^5,$$

其中 $\xi_1, \xi_2, \xi_3 \in [a,b]$, 而函数 f 属于相应的函数类 $C^{(k)}[a,b]$.

d) 设 f 是多项式 P. 为使矩形公式、梯形公式和抛物形公式 是准确的, 问多项式 P 的最高次数分别是多少?

设 $h = \frac{b-a}{n}$; $x_k = a + hk, k = 0, 1, \cdots, n$; $y_k = f(x_k)$.

e) 试证: 在矩形公式

$$\int_a^b f(x)dx = h(y_0 + y_1 + \cdots + y_n) + R_1$$

中, 余项 $R_1 = \frac{f'(\xi)}{2}(b-a)h$, 其中 $\xi \in [a,b]$.

---

① 辛普森(Simpson) (1710—1761)—— 英国数学家.

f) 试证: 在梯形公式

$$\int_a^b f(x)dx = \frac{h}{2}[(y_0 + y_n) + 2(y_1 + y_2 + \cdots + y_{n-1})] + R_2$$

中, 余项 $R_2 = -\frac{f^{(2)}(\xi)}{12}(b-a)h^2$, 其中 $\xi \in [a,b]$.

g) 试证: 在辛普森公式 (抛物形公式)

$$\int_a^b f(x)dx = \frac{h}{3}[(y_0 + y_n) + 4(y_1 + y_3 + \cdots + y_{n-1}) + 2(y_2 + y_4 + \cdots + y_{n-2})] + R_3$$

中, 这里 n 是偶数, $R_3 = -\frac{f^{(4)}(\xi)}{180}(b-a)h^4$, 其中 $\xi \in [a,b]$.

h) 试根据关系

$$\pi = 4\int_0^1 \frac{dx}{1+x^2}$$

计算 π的值, 要求精确到 10^{-3}, 依次利用矩形公式、梯形公式和抛物形公式. 注意辛普森公式是特别有效的. 正因为如此, 它是最常用的求积公式 (对一维情形下数值积分公式采用这些命名, 是因为它们可把积分与相应的曲边梯形的面积等同起来).

10. 变换公式 (7), 求出如下形式的泰勒公式余项, 其中 $h = x - a$:

a) $\frac{h^n}{(n-1)!}\int_0^1 f^{(n)}(a+\tau h)(1-\tau)^{n-1}d\tau$;

b) $\frac{h^n}{n!}\int_0^1 f^{(n)}(x - h\sqrt[n]{t})dt$.

11. 试证: 积分中的重要变量替换公式 (9), 在没有替换函数单调的假定时, 仍然成立.

§4. 积分的一些应用

在应用中经常是按照同一个格式运用积分, 因此, 单独地叙述一下这个格式是有益的. 这就是本节第一段的内容.

1. 定向区间的可加函数和积分 在 §2 中讨论积分的可加性时, 我们引进了定向区间的可加函数的概念. 我们记得, 定向区间的可加函数是形如 $(\alpha,\beta) \mapsto I(\alpha,\beta)$ 的函数, 它使由固定区间 $[a,b]$ 的任意两点 $\alpha,\beta \in [a,b]$ 构成的每个序对 (α,β) 对应于一个数 $I(\alpha,\beta)$, 而且对于任意三点 $\alpha,\beta,\gamma \in [a,b]$ 成立等式

$$I(\alpha,\gamma) = I(\alpha,\beta) + I(\beta,\gamma). \tag{1}$$

当 $\alpha = \beta = \gamma$ 时, 从 (1) 推出 $I(\alpha,\alpha) = 0$; 而当 $\alpha = \gamma$ 时, 我们得到 $I(\alpha,\beta) + I(\beta,\alpha) = 0$, 亦即 $I(\alpha,\beta) = -I(\beta,\alpha)$. 在这里可以看出点 α,β 的次序的影响.

令

$$\mathcal{F}(x) = I(a,x),$$

根据函数 I 的可加性, 我们有

$$I(\alpha,\beta)=I(a,\beta)-I(a,\alpha)=\mathcal{F}(\beta)-\mathcal{F}(\alpha).$$

这样一来, 每个定向区间的可加函数都有

$$I(\alpha,\beta)=\mathcal{F}(\beta)-\mathcal{F}(\alpha) \tag{2}$$

的形式, 其中 $x\mapsto\mathcal{F}(x)$ 是区间 $[a,b]$ 上点的函数.

容易验证, 上述断言的逆命题也对, 亦即: 定义在区间 $[a,b]$ 上的任意一个函数 $x\mapsto\mathcal{F}(x)$, 按照公式 (2) 都产生一个定向区间的可加函数.

我们举两个典型例子

例 1 如果 $f\in\mathcal{R}[a,b]$, 则函数 $\mathcal{F}(x)=\int_a^x f(t)dt$ 按公式 (2) 产生可加函数

$$I(\alpha,\beta)=\int_\alpha^\beta f(t)dt.$$

我们注意到, 在这种情形下, 函数 $\mathcal{F}(x)$ 在区间 $[a,b]$ 上连续.

例 2 设区间 $[0,1]$ 是无重量的一条弦, 在 $x=1/2$ 处系一个质量为 1 的珠子. 设 $\mathcal{F}(x)$ 是位于区间 $[0,x]$ 一段弦的质量, 则由假设条件, 有

$$\mathcal{F}(x)=\begin{cases}0, 当 x<\frac{1}{2},\\ 1, 当 \frac{1}{2}\leqslant x\leqslant 1.\end{cases}$$

可加函数

$$I(\alpha,\beta)=\mathcal{F}(\beta)-\mathcal{F}(\alpha)$$

当 $\beta>\alpha$ 的物理意义是半开区间 $]\alpha,\beta]$ 中的质量.

由于函数 $\mathcal{F}$ 间断, 在这种情形可加函数 $I(\alpha,\beta)$ 不可能表示成某个函数 —— 质量密度 —— 的黎曼积分. (这个密度, 亦即区间的质量与区间长之比, 在区间 $[0,1]$ 上任意点 $x\neq\frac{1}{2}$ 处应当等于 0, 而在 $x=\frac{1}{2}$ 处应当是无穷大.)

现在来证明一个以后有用的命题, 它给出了可加函数能由积分产生的一个充分条件.

命题 1 *设可加函数 $I(\alpha,\beta)$ 对于区间 $[a,b]$ 的点 α,β 有定义, 且存在函数 $f\in\mathcal{R}[a,b]$, 它与 I 以下述方式相联系: 对于任意区间 $[\alpha,\beta], a\leqslant\alpha<\beta\leqslant b$, 成立关系*

$$\inf_{x\in[\alpha,\beta]}f(x)(\beta-\alpha)\leqslant I(\alpha,\beta)\leqslant\sup_{x\in[\alpha,\beta]}f(x)(\beta-\alpha).$$

那么,

$$I(a,b)=\int_a^b f(x)dx.$$

◀ 设 P 是区间 $[a,b]$ 的任意一个分划:$a=x_0<x_1<\cdots<x_n=b$;

$$m_i=\inf_{x\in[x_{i-1},x_i]} f(x),\quad M_i=\sup_{x\in[x_{i-1},x_i]} f(x).$$

对于分划 P 的每个区间 $[x_{i-1},x_i]$, 根据假设条件, 我们有

$$m_i\Delta x_i\leqslant I(x_{i-1},x_i)\leqslant M_i\Delta x_i.$$

对于这个不等式求和并利用函数 $I(\alpha,\beta)$ 的可加性, 得到

$$\sum_{i=1}^n m_i\Delta x_i\leqslant I(a,b)\leqslant\sum_{i=1}^n M_i\Delta x_i.$$

我们知道, 这个不等式两头分别是函数 f 相应于区间 $[a,b]$ 的分划 P 的下积分和与上积分和. 当 $\lambda(P)\to 0$ 时, 它们有同样的极限, 即 f 在 $[a,b]$ 上的积分. 这样一来, 取极限 $\lambda(P)\to 0$, 就得到

$$I(a,b)=\int_a^b f(x)dx. ▶$$

现在, 我们以具体问题来说明命题 1 的作用.

2. 道路的长度 有一质点在空间 $\mathbb{R}^3$ 中运动. 设已知它的运动规律是 $r(t)=(x(t),y(t),z(t))$, 其中 $x(t),y(t),z(t)$ 是点在时刻 t 的笛卡儿直角坐标.

我们要确定在时间间隔 $a\leqslant t\leqslant b$ 内质点所通过的道路的长度 $l[a,b]$.

首先把几个概念精确化.

定义 1 设 $x(t),y(t),z(t)$ 是定义在区间上的连续函数, 则从这个区间到空间 $\mathbb{R}^3$ 的映射 $t\mapsto(x(t),y(t),z(t))$ 叫做 $\mathbb{R}^3$ 中的一条*道路*.

定义 2 如果 $t\mapsto(x(t),y(t),z(t))$ 是一条道路. 它的参数 t 的变化范围是闭区间 $[a,b]$, 则称空间 $\mathbb{R}^3$ 中的点 $A=(x(a),y(a),z(a)),B=(x(b),y(b),z(b))$ 分别是该道路的*起点*和*终点*.

定义 3 一条道路叫做*闭的*, 如果它有起点和终点而且它们是重合的.

定义 4 如果 $\Gamma:I\to\mathbb{R}^3$ 是一条道路, 则区间 I 在空间 $\mathbb{R}^3$ 的像 $\Gamma(I)$ 叫做该道路的*承载子*.

抽象道路的承载子完全可能不像我们称之为曲线的几何图形. 譬如, 有这样的道路, 其承载子包含整个三维立方体 (佩亚诺 "曲线"). 但是, 如果函数 $x(t),y(t),z(t)$

充分正规 (譬如, 机械运动的情形, 这时它们是可微的), 则可严格地证明它没有任何与我们的直观相抵触的东西.

定义 5 设 $\Gamma : I \to \mathbb{R}^3$. 如果映射 $I \to \Gamma(I)$ 是双方单值的, 则 Γ 叫做*简单道路*或*参数化曲线*, 其承载子叫做 $\mathbb{R}^3$ 中的*曲线*.

定义 6 闭道路 $\Gamma : [a, b] \to \mathbb{R}^3$ 叫做*简单闭道路*或*简单闭曲线*, 如果道路 $\Gamma : [a, b] \to \mathbb{R}^3$ 是简单的.

这就是说, 简单道路与任意道路的区别在于当我们沿着它的承载子运动时不会回到曾经经过的任何点, 亦即在任何地方它都不与自己的轨迹相交, 除非是在终点 (此时, 给定的简单道路是闭的).

定义 7 设 $\Gamma : I \to \mathbb{R}^3$. 称它是*具某种光滑性的某某类的道路*, 如果确定这个道路的函数 $x(t), y(t), z(t)$ 属于相应的函数类.

例如, 我们有 $C[a, b]$ 类, $C^{(1)}[a, b]$ 类以及 $C^{(k)}[a, b]$ 类道路.

定义 8 设 $\Gamma : [a, b] \to \mathbb{R}^3$. 称它是*分段光滑的*, 如果能把区间 $[a, b]$ 分成有限多个区间, 且在其中每一个上, 映射 Γ 都由连续可微函数确定.

光滑道路, 也就是 $C^{(1)}$ 类道路, 以及分段光滑道路正是我们现在要研究的对象.

现在可以把原来的问题陈述成如下形式: 试确定一光滑道路 $\Gamma : [a, b] \to \mathbb{R}^3$ 的长度.

关于在时间间隔 $\alpha \leqslant t \leqslant \beta$ 内所通过的道路的长度 $l[a, b]$, 我们原来的概念是这样的: 第一, 如果 $\alpha < \beta < \gamma$, 则

$$l[\alpha, \gamma] = l[\alpha, \beta] + l[\beta, \gamma];$$

第二, 如果 $\boldsymbol{v}(t) = (\dot{x}(t), \dot{y}(t), \dot{z}(t))$ 是点在时刻 t 的速度, 则

$$\inf_{t \in [\alpha, \beta]} |\boldsymbol{v}(t)|(\beta - \alpha) \leqslant l[\alpha, \beta] \leqslant \sup_{t \in [\alpha, \beta]} |\boldsymbol{v}(t)|(\beta - \alpha).$$

这样一来, 如果函数 $\dot{x}(t), \dot{y}(t), \dot{z}(t)$ 在 $[a, b]$ 上连续可微, 则根据命题 1, 我们能定义

$$l[a, b] = \int_a^b |\boldsymbol{v}(t)| dt = \int_a^b \sqrt{\dot{x}^2(t) + \dot{y}^2(t) + \dot{z}^2(t)} dt, \tag{3}$$

我们就取它作为光滑道路 $\Gamma : [a, b] \to \mathbb{R}^3$ 的长度的定义.

如果 $z(t) = 0$, 则道路的承载子在一个平面内, 而公式 (3) 有如下形式:

$$l[a, b] = \int_a^b \sqrt{\dot{x}^2(t) + \dot{y}^2(t)} dt. \tag{4}$$

例 3 我们做一些把公式 (4) 应用于已知对象的尝试. 设有一点按规律

$$x = R\cos 2\pi t,$$
$$y = R\sin 2\pi t, \tag{5}$$

在平面内运动.

在时间间隔 $[0,1]$ 内该点绕半径为 R 的圆周一次, 通过的路程长应当是 $2\pi R$. 根据公式 (4) 计算可得:

$$l[a,b] = \int_0^1 \sqrt{(-2\pi R\sin 2\pi t)^2 + (2\pi R\cos 2\pi t)^2}dt = 2\pi R.$$

尽管这两个结果的一致是令人鼓舞的, 但值得注意的是在上面的推理中存在一些逻辑上的漏洞.

如果采用中学里的定义, 函数 $\cos\alpha$ 和 $\sin\alpha$ 是点 $p_0 = (1,0)$ 旋转 α 角所得的像 p 的笛卡儿坐标.

不计符号, α 的值是以圆周 $x^2 + y^2 = 1$ 上界于 p_0 和 p 之间的弧的长来度量的. 这样一来, 在这种处理下, 三角函数 的定义以圆弧的长的概念为基础. 因此, 当计算上面的圆周长时, 由于用了圆周的参数方程 (5), 从一定意义上讲, 是陷入了逻辑循环.

但是, 我们将会看到, 这个困难并不是原则性的, 因为圆周的参数形式完全可以不用三角函数给出.

考察定义在区间 $[a,b] \subset \mathbb{R}$ 上的函数 $y = f(x)$ 的图形的长. 这是计算具有特殊参数化形式 $x \mapsto (x, f(x))$ 的道路 $\Gamma : [a,b] \to \mathbb{R}^2$ 的长度的问题. 从这个参数化形式看出, 映射 $\Gamma : [a,b] \to \mathbb{R}^2$ 是双方单值的. 因此, 根据定义 5, 函数的图形是 $\mathbb{R}^2$ 中的曲线.

这时, 公式 (4) 可以简化, 因为在其中令 $x = t, y = f(t)$ 就得到

$$l[a,b] = \int_a^b \sqrt{1 + [f'(x)]^2}dx. \tag{6}$$

特别地, 如果考虑圆周 $x^2 + y^2 = 1$ 的半周圆 $y = \sqrt{1-x^2}, -1 \leqslant x \leqslant 1$, 则得到

$$l = \int_{-1}^1 \sqrt{1 + \left[\frac{-x}{\sqrt{1-x^2}}\right]^2}dx = \int_{-1}^1 \frac{dx}{\sqrt{1-x^2}} \tag{7}$$

但是, 在最后的积分中被积函数是无界的, 因此, 在前述意义下它是不存在的. 这难道说半圆周没有长度吗? 这仅仅说明半圆周的上述参数化不满足使函数 $\dot{x}, \dot{y}$ 连续的条件, 而这些条件能保证公式 (4), 从而也保证公式 (6) 成立. 因此, 我们应当或者考虑推广积分的概念, 使 (7) 式中的积分有确定的意义; 或者选取另外的参数化形式使之能满足保证公式 (6) 可用的条件.

我们看到, 如果在形如 $[-1+\delta,1-\delta]$ (其中 δ 满足条件 $-1<-1+\delta<1-\delta<1$) 的区间上考察上面所取的参数形式, 则公式 (6) 可用, 根据它求出在区间 $[-1+\delta,1-\delta]$ 上的圆弧的长为

$$l[-1+\delta,1-\delta]=\int_{-1+\delta}^{1-\delta}\frac{dx}{\sqrt{1-x^2}}.$$

因此, 自然应当认为, 半圆周的长 l 是极限 $\lim\limits_{\delta\to+0}l[-1+\delta,1-\delta]$. 正是在这个意义下理解关系 (7) 中的积分. 下一节我们将详细讨论黎曼积分概念的这种自然的推广. 至于说到我们研究的具体问题, 那么, 既使不改换参数形式也可以求出, 譬如, 长度 $l\left[-\frac{1}{2},\frac{1}{2}\right]$, 它是单位圆圆周上所对的弦等于半径的那样一段弧的长度. 这样,(从几何知识可知) 应当成立: $l=3\cdot l\left[-\frac{1}{2},\frac{1}{2}\right]$.

我们还看到

$$\int\frac{dx}{\sqrt{1-x^2}}=\int\frac{(1-x^2+x^2)dx}{\sqrt{1-x^2}}=\int\sqrt{1-x^2}dx-\frac{1}{2}\int\frac{xd(1-x^2)}{\sqrt{1-x^2}}$$
$$=2\int\sqrt{1-x^2}dx-x\sqrt{1-x^2},$$

因此,

$$l[-1+\delta,1-\delta]=2\int_{-1+\delta}^{1-\delta}\sqrt{1-x^2}dx-(x\sqrt{1-x^2})\Big|_{-1+\delta}^{1-\delta}.$$

这样一来

$$l=\lim_{\delta\to+0}l[-1+\delta,1-\delta]=2\int_{-1}^{1}\sqrt{1-x^2}dx.$$

半径为 1 的半圆周的长度以符号 π表示, 于是, 我们得到公式:

$$\pi=2\int_{-1}^{1}\sqrt{1-x^2}dx.$$

最后这个积分是常义 (不是广义的) 黎曼积分, 从而可以计算出它的具有任意给定的精确度的值.

如果对于 $x\in[-1,1]$ 称 $l[x,1]$ 为 $\arccos x$, 则根据上面的计算, 有

$$\arccos x=\int_x^1\frac{dx}{\sqrt{1-x^2}}$$

或

$$\arccos x=x\sqrt{1-x^2}+2\int_x^1\sqrt{1-x^2}dx.$$

如果把弧长作为原始的概念, 那么, 刚才引入的函数 $x \mapsto \arccos x$, 以及可用类似方法引入函数 $x \mapsto \arcsin x$ 都应作为原始概念, 而函数 $x \mapsto \cos x$ 和 $x \mapsto \sin x$ 可以作为它们在相应区间上的反函数得到. 实际上, 在初等几何中正是这样做的.

计算半圆周长这个例子的教益不仅在于通过对它的分析使我们有可能对三角函数的定义作出某些有益的评注, 还在于它提出了一个很自然的问题: 当论及曲线长度时, 由公式 (3) 确定的数是否依赖于 x、y、z 坐标系和曲线的参数形式的选取?

我们把对空间笛卡儿坐标的作用的分析留给读者自己去做, 这里只讨论参数形式的作用.

准确地说, 设 $\Gamma : I \to \mathbb{R}^3$ 是一条简单道路, 它的承载子是一条曲线. 给定了一条道路, 就是建立了相应曲线的*参数形式*. 点或数 $t \in I$ 叫做*参数*, 而区间 I 叫做参数的*变化区域*.

如果 $\Gamma : I \to \mathcal{L}$ 和 $\tilde{\Gamma} : \tilde{I} \to \mathcal{L}$ 是具有相同值域 $\mathcal{L}$ 的两个双方单值映射, 则显然可产生这两个映射的定义域间的两个双方单值映射 $\tilde{\Gamma}^{-1} \circ \Gamma : I \to \tilde{I}$ 和 $\Gamma^{-1} \circ \tilde{\Gamma} : \tilde{I} \to I$.

特别地, 如果一条曲线有两个参数形式, 则在它们的参数 $t \in I$ 和 $\tau \in \tilde{I}$ 之间可以建立自然的对应 $t = t(\tau)$ 或 $\tau = \tau(t)$, 使得能根据点在一个参数形式下的参数值确定它在另一参数形式下的参数值.

设 $\Gamma : [a, b] \to \mathcal{L}, \tilde{\Gamma} : [\alpha, \beta] \to \mathcal{L}$ 是同一条曲线的两个参数形式, 且在起点和终点分别有 $\Gamma(a) = \tilde{\Gamma}(\alpha)$ 和 $\Gamma(b) = \tilde{\Gamma}(\beta)$. 那么, 从一个参数变到另一个参数的变换函数 $t = t(\tau), \tau = \tau(t)$ 是把 $a \leqslant t \leqslant b, \alpha \leqslant \tau \leqslant \beta$ 这两个区间中的一个映成另一个的连续、单调映射, 而且 $a \leftrightarrow \alpha$ 和 $b \leftrightarrow \beta$. 如果曲线 Γ 和 $\tilde{\Gamma}$ 分别由光滑函数组 $(x(t), y(t), z(t))$ 和 $(\tilde{x}(\tau), \tilde{y}(\tau), \tilde{z}(\tau))$ 给出, 且 $|v(t)|^2 = \dot{x}^2(t) + \dot{y}^2(t) + \dot{z}^2(t) \neq 0$ 在 $[a, b]$ 上成立, 而在 $[\alpha, \beta]$ 上有 $|v(t)|^2 = \dot{\tilde{x}}^2(t) + \dot{\tilde{y}}^2(t) + \dot{\tilde{z}}^2(t) \neq 0$. 那么, 可以验证, 这时变换函数 $t = t(\tau)$ 和 $\tau = \tau(t)$ 是光滑的, 而且在各自的定义区间上有正导数.

我们现在不证明这个命题, 它将在适当的时候作为隐函数定理的一个推论而得到. 而现时, 上述命题促使我们引出以下定义.

定义 9 称道路 $\tilde{\Gamma} : [\alpha, \beta] \to \mathbb{R}^3$ 是从道路 $\Gamma : [a, b] \to \mathbb{R}^3$ *借助于参数的容许替换得出的*, 如果存在光滑映射 $T : [\alpha, \beta] \to [a, b]$ 使 $T(\alpha) = a, T(\beta) = b$, 在 $[\alpha, \beta]$ 上有 $T'(\tau) > 0$, 且 $\tilde{\Gamma} = \Gamma \circ T$.

现在来证明如下一般的命题.

命题 2 *如果光滑道路 $\tilde{\Gamma} : [\alpha, \beta] \to \mathbb{R}^3$ 是从光滑道路 $\Gamma : [a, b] \to \mathbb{R}^3$ 借助于参数的容许替换得出的, 则这两条道路的长相等.*

◀ 设 $\tilde{\Gamma} : [\alpha, \beta] \to \mathbb{R}^3$ 和 $\Gamma : [a, b] \to \mathbb{R}^3$ 分别由光滑函数组 $\tau \mapsto (\tilde{x}(\tau), \tilde{y}(\tau), \tilde{z}(\tau))$ 和 $t \mapsto (x(t), y(t), z(t))$ 给出, 而 $t = t(\tau)$ 是参数的容许替换, 在这个变换下有 $\tilde{x}(\tau) = x(t(\tau)), \tilde{y}(\tau) = y(t(\tau)), \tilde{z}(\tau) = z(t(\tau))$.

利用定义 3, 复合函数微分法以及积分变量替换规则, 得到

$$\begin{aligned}\int_a^b \sqrt{\dot{x}^2(t)+\dot{y}^2(t)+\dot{z}^2(t)}dt &= \int_\alpha^\beta \sqrt{\dot{x}^2(t(\tau))+\dot{y}^2(t(\tau))+\dot{z}^2(t(\tau))}t'(\tau)d\tau \\ &= \int_\alpha^\beta \sqrt{[\dot{x}(t(\tau))t'(\tau)]^2+[\dot{y}(t(\tau))t'(\tau)]^2+[\dot{z}(t(\tau))t'(\tau)]^2}d\tau \\ &= \int_\alpha^\beta \sqrt{\dot{\tilde{x}}^2(\tau)+\dot{\tilde{y}}^2(\tau)+\dot{\tilde{z}}^2(\tau)}d\tau. \blacktriangleright\end{aligned}$$

特别地, 我们也证明了曲线的长不依赖于它的光滑参数形式的选取.

分段光滑道路的长定义为组成它的诸光滑道路长之和. 因此, 容易验证, 分段光滑道路的长经过参数的容许替换仍然不变.

作为结束关于道路的长和曲线的长这些概念的讨论 (在得到命题 1 之后我们就有权正式使用这些概念了), 我们再研究一个例子.

例 4 求用标准方程

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1 \tag{8}$$

(其中 $a \geqslant b > 0$) 给出的椭圆的周长.

取参数形式 $x = a\sin\psi, y = b\cos\psi, 0 \leqslant \psi \leqslant 2\pi$. 我们得到

$$\begin{aligned}l &= \int_0^{2\pi} \sqrt{(a\cos\psi)^2+(-b\sin\psi)^2}d\psi \\ &= \int_0^{2\pi} \sqrt{a^2-(a^2-b^2)\sin^2\psi}d\psi \\ &= 4a\int_0^{\pi/2} \sqrt{1-\frac{a^2-b^2}{a^2}\sin^2\psi}d\psi \\ &= 4a\int_0^{\pi/2} \sqrt{1-k^2\sin^2\psi}d\psi,\end{aligned}$$

其中 $k^2 = 1-\dfrac{b^2}{a^2}$ 是椭圆的离心率的平方.

积分

$$E(k,\varphi) = \int_0^\varphi \sqrt{1-k^2\sin^2\psi}d\psi$$

不能表示成初等函数, 而且, 由于它与椭圆有上述关系, 所以叫做*椭圆积分*. 更准确地说, $E(k,\varphi)$ 叫*第二类勒让德椭圆积分*. 当 $\varphi = \pi/2$ 时, 这个积分的值, 显然它只依赖于 k, 记做 $E(k)$, 叫*第二类全椭圆积分*. 因此, $E(k) = E(k,\pi/2)$, 而椭圆的周长 $l = 4aE(k)$.

3. 曲边梯形的面积 我们来研究图形 $aABb$(图 48), 它叫做*曲边梯形*. 这个图形由垂直线段 aA, bB, 横轴上的线段 $[a,b]$ 和曲线 $\breve{AB}$ 围成, 其中 $\breve{AB}$ 是区间 $[a,b]$ 上的一个可积函数 $y=f(x)$ 的图像.

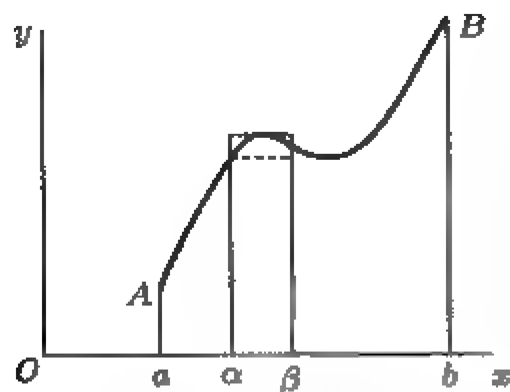


图 48

设 $[\alpha,\beta]$ 是包含在 $[a,b]$ 中的一个区间. 以 $S(\alpha,\beta)$ 表示与 $[\alpha,\beta]$ 相应的曲边梯形 $\alpha f(\alpha)f(\beta)\beta$ 的面积.

我们关于面积的概念是这样的: 如果 $a \leqslant \alpha < \beta < \gamma \leqslant b$, 则

$$S(\alpha,\gamma) = S(\alpha,\beta) + S(\beta,\gamma)$$

(面积的可加性) 以及

$$\inf_{x\in[\alpha,\beta]} f(x)(\beta-\alpha) \leqslant S(\alpha,\beta) \leqslant \sup_{x\in[\alpha,\beta]} f(x)(\beta-\alpha)$$

(一个图形的面积不小于它所包含的图形的面积).

因此, 根据命题 1, 上述图形的面积应按公式

$$S(a,b) = \int_a^b f(x)dx \tag{9}$$

计算.

例 5 用公式 (9) 计算由标准方程 (8) 给出的椭圆的面积.

根据图形的对称性和面积的可加性, 只要求出该椭圆位于第一象限的那一部分的面积然后将它四倍即可. 我们来进行计算:

$$S = 4\int_0^a \sqrt{b^2\left(1-\frac{x^2}{a^2}\right)}dx = 4b\int_0^{\pi/2} \sqrt{1-\sin^2 t}a\cos t dt$$
$$= 4ab\int_0^{\pi/2} \cos^2 t dt = 2ab\int_0^{\pi/2} (1-\cos 2t)dt = \pi ab.$$

在这个计算过程中我们作了变换 $x = a\sin t, 0 \leqslant t \leqslant \pi/2$.

因此, $S = \pi ab$. 特别地, 当 $a = b = R$ 时, 我们得到半径为 R 的圆的面积公式 πR^2.

注 必须指出, 公式 (9) 是在条件 "在 $[a,b]$ 上 $f(x) \geqslant 0$" 下给出曲边梯形面积的. 而如果是任意一个可积函数, 则积分 (9) 给出的显然是位于横轴上边和横轴下边的那些曲边梯形的面积 的代数和. 这里在横轴上边的曲边梯形的面积带有正号, 而横轴下方的则带负号.

4. 旋转体的体积　设图 48 上的曲边梯形绕区间 $[a,b]$ 旋转, 我们来确定所得立体的体积.

以 $V(\alpha,\beta)$ 表示由曲边梯形$\alpha f(\alpha)f(\beta)\beta$ 旋转所得立体 (见图 48) 的体积, 其中相应的区间 $[\alpha,\beta] \subset [a,b]$.

根据我们关于体积的概念, 应当成立以下关系: 如果 $a \leqslant \alpha < \beta < \gamma \leqslant b$, 则

$$V(\alpha,\gamma) = V(\alpha,\beta) + V(\beta,\gamma)$$

和

$$\pi(\inf_{x\in[\alpha,\beta]} f(x))^2(\beta-\alpha) \leqslant V(\alpha,\beta) \leqslant \pi(\sup_{x\in[\alpha,\beta]} f(x))^2(\beta-\alpha).$$

在上式中, 我们用了内接和外切柱体的体积去估计 $V(\alpha,\beta)$, 还用了柱体体积公式 (如果已经知道了圆面积公式就不难得到它).

因此, 根据命题 1, 有

$$V(a,b)=\pi\int_a^b f^2(x)dx. \tag{10}$$

例 6 将由横轴上的区间 $[-R,R]$ 和圆弧 $y=\sqrt{R^2-x^2}(-R\leqslant x\leqslant R)$ 所围成的半圆绕横轴旋转, 就得到半径为 R 的三维球, 根据公式 (10) 容易算出它的体积

$$V=\pi\int_{-R}^{R}(R^2-x^2)dx=\frac{4}{3}\pi R^3.$$

对线、面以及体的度量的更详尽的讨论将在本教程第 II 卷中进行, 那时我们还将解决给出的定义的不变性问题.

5. 功与能 物体在常力作用下沿着力的作用方向移动. 力对物体所做的功以力和位移的乘积 $F\cdot S$ 来度量. 这个量叫做力在给定位移上的*功*. 在一般情形, 给定力的方向与位移方向可能不共线 (譬如用绳子拖雪撬), 这时功是用力向量与位移向量的内积 $\langle F,S\rangle$ 来度量.

下面几个例子涉及功的计算并利用了与功相联系的能量的概念.

例 7 根据定义, 把质量为 m 的物体从离地面高度为 h_1 的地方举到高度为 h_2 的地方, 克服重力所做的功等于 $mg(h_2-h_1)$. 假定整个操作过程是在地面附近进行的, 从而重力 mg 的变化可以忽略不计. 关于一般情形将在例 10 中讨论.

例 8* 设有一理想的弹簧, 它的一端固定在负半轴上一点, 且使数轴的 O 点与另一端的平衡位置重合. 众所周知, 当这个端点位于 x 点时, 拉住它所需要的力等于 kx, 此处 k 是弹簧的弹性系数.

现在来计算将弹簧的右端点从 $x=a$ 处移动到 $x=b$ 处所需之功.

我们认为功 $A(\alpha,\beta)$ 是区间 $[\alpha,\beta]$ 的可加函数并满足估计

$$\inf_{x\in[\alpha,\beta]} kx(\beta-\alpha) \leqslant A(\alpha,\beta) \leqslant \sup_{x\in[\alpha,\beta]} kx(\beta-\alpha),$$

根据命题 1, 我们得到

$$A(a,b)=\int_a^b kxdx=\frac{kx^2}{2}\Big|_a^b.$$

*原著中该例的叙述有误, 这里作了修改. 译者注.

这是克服弹性力所需之功, 而弹性力本身在同一段位移上做的功与此只差一个符号.

我们求出来的函数 $U(x)=\dfrac{kx^2}{2}$ 可以用来计算改变弹簧状态时所要完成的功, 而这意味着, 当弹簧恢复到原来的状态时它将完成这些功. 这种只依赖于系统构形的函数叫做该系统的势能. 从势能的构成看出, 它的导数给出了弹性力, 但具有与弹性力相反的符号.

如果质点 m 在上述弹性力的作用下沿 x 轴运动, 则它的坐标 $x(t)$ 作为时间的函数将满足方程

$$m\ddot{x}=-kx. \tag{11}$$

有一次我们验证了 (见第五章 §6 第 6 段) 量

$$\frac{mv^2}{2}+\frac{kx^2}{2}=K(t)+U(x(t))=E \tag{12}$$

(即系统的动能和势能之和) 在运动过程中保持不变.

例 9 现在再考察一个例子, 在这个例子中将遇到一系列我们已经掌握的微分学和积分学的概念.

首先注意, 类似于为描述满足方程 (11) 的具体力学系统而引进的函数 (12), 对于形如

$$\ddot{s}(t)=f(s(t)) \tag{13}$$

的任意方程, 其中 $f(s)$ 是给定的函数, 只要取 $U(s)$ 满足 $U'(s)=-f(s)$, 就能得到和式

$$\frac{\dot{s}^2}{2}+U(s)=E \tag{14}$$

不随时间变化.

事实上,

$$\frac{dE}{dt}=\frac{1}{2}\frac{d\dot{s}^2}{dt}+\frac{dU(s)}{dt}=\dot{s}\ddot{s}+\frac{dU}{ds}\frac{ds}{dt}=\dot{s}(\ddot{s}-f(s))=0.$$

这样一来, 从 (14) 求出 $\dot{s}=\pm\sqrt{2(E-U(s))}$(这里根式的符号应与导数 $\dfrac{ds}{dt}$ 的符号一致), 从而

$$\frac{dt}{ds}=\pm\frac{1}{\sqrt{2(E-U(s))}}.$$

由此,

$$t=c_1\pm\int\frac{ds}{\sqrt{2(E-U(s))}},$$

其中 E 是常数.



图 49

因此, 利用方程 (13) 的能量守恒律 (14), 原则上可以解出这个方程, 但求出的不是函数 $s(t)$, 而是它的反函数 $t(s)$.

方程 (13) 出现于, 譬如, 描述点沿着给定曲线运动的问题中. 设质点在重力作用下沿着一个狭窄的理想光滑槽子移动 (图 49). 设 $s(t)$ 是从某固定点 0——计算的起点, 沿着槽子到时刻 t 质点所处位置的距离 (即路程的长). 显然, $\dot{s}(t)$ 是质点的速度, 而 $\ddot{s}(t)$ 是它的加速度切向分量, 它等于重力在小槽的给定点处的切线分量. 重力的切线分量只依赖于小槽的点, 也就是说它只依赖于 s (因为 s 可以看作小槽所在曲线的参数[①]), 这是很明显的. 如果以 $f(s)$ 表示重力的这个分力, 则

$$m\ddot{s} = f(s)$$

对于这个方程, 有

$$\frac{1}{2}m\dot{s}^2 + U(s) = E$$

保持不变, 其中 $U(s)$ 满足 $U'(s) = -f(s)$.

由于 $\frac{1}{2}m\dot{s}^2$ 这一项是点的动能, 而沿槽子运动假定没有摩擦, 因此, 无须计算就可猜到, 不计常数之差, 函数 $U(s)$ 应具有 $mgh(s)$ 的形式, 这里 $mgh(s)$ 是重力场中位于高度 $h(s)$ 处的点的势能.

如果在初始时刻 $t=0$ 有 $\dot{s}(0)=0, s(0)=s_0$ 和 $h(s_0)=h_0$, 则从关系

$$\frac{2E}{m} = \dot{s}^2 + 2gh(s) = C$$

可以求出, $C = 2gh_0$. 因此, $\dot{s}^2 = 2g(h_0 - h(s))$,

$$t = \int_{s_0}^{s} \frac{ds}{\sqrt{2g(h_0 - h(s))}}. \tag{15}$$

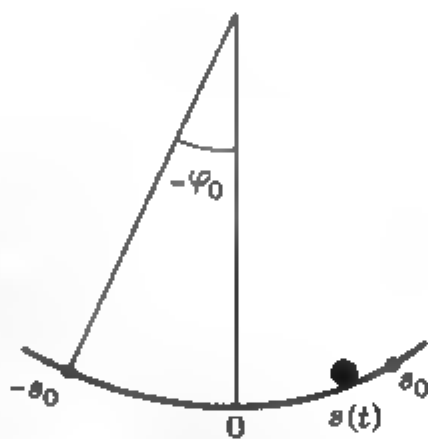


图 50

特别地, 如果点沿着半径为 R 的圆周运动 (譬如摆的振动), 长度 s 从圆周的最低点算起, 而初始条件如下: 当 $t=0$ 时 $\dot{s}(0)=0$ 并给出了初始倾角 φ_0(图 50), 那么, 容易验证, 以倾角 φ 表示 s 和 $h(s)$ 后, 有

$$t = \int_{-s_0}^{s} \frac{ds}{\sqrt{2g(h_0 - h(s))}} = \int_{-\varphi_0}^{\varphi} \frac{R\,d\psi}{\sqrt{2gR(\cos\psi - \cos\varphi_0)}},$$

或

$$t = \frac{1}{2}\sqrt{\frac{R}{g}} \int_{-\varphi_0}^{\varphi} \frac{d\psi}{\sqrt{\sin^2\frac{\varphi_0}{2} - \sin^2\frac{\psi}{2}}}. \tag{16}$$

①借助于曲线的长对曲线做的参数化叫做自然参数化, 这时 s 叫自然参数.

这样一来, 我们得到摆的半周期

$$\frac{1}{2}T=\frac{1}{2}\sqrt{\frac{R}{g}}\int_{-\varphi_0}^{\varphi_0}\frac{d\psi}{\sqrt{\sin^2\frac{\varphi_0}{2}-\sin^2\frac{\psi}{2}}},\tag{17}$$

由此, 经过代换

$$\frac{\sin\frac{\psi}{2}}{\sin\frac{\varphi_0}{2}}=\sin\theta,$$

我们得到

$$T=4\sqrt{\frac{R}{g}}\int_0^{\frac{\pi}{2}}\frac{d\theta}{\sqrt{1-k^2\sin^2\theta}},\tag{18}$$

其中 $k^2=\sin^2\frac{\varphi_0}{2}$.

我们记得, 函数

$$F(k,\varphi)=\int_0^{\varphi}\frac{d\theta}{\sqrt{1-k^2\sin^2\theta}}$$

叫做第一类勒让德椭圆积分. 当 $\varphi=\frac{\pi}{2}$ 时, 它只依赖于 k^2, 记做 $K(k)$, 叫做第一类全椭圆积分. 这样, 摆的振动周期等于

$$T=4\sqrt{\frac{R}{g}}K(k).\tag{19}$$

如果初始倾角很小, 则可令 $k=0$, 从而得到近似公式

$$T\approx 2\pi\sqrt{\frac{R}{g}}.\tag{20}$$

在得到公式 (18) 后, 我们应当分析一下推导的全过程. 我们将发现,(15)—(17) 中被积函数在积分区间上是无界的. 类似的困难在讨论曲线长的时候就遇到过. 因此, 我们大体上已知道应当赋予积分 (15)—(17) 怎样的意义.

但是, 一次又一次的产生这个问题, 使我们认识到有必要仔细地研究一下它的精确数学提法. 这些将在下一节去做.

例 10 在地面上空将质量为 m 的物体沿着轨线 $t\mapsto(x(t),y(t),z(t))$ 举起, 这里 t 是时间, $a\leqslant t\leqslant b$, 而 x,y,z 是点在空间中的笛卡儿坐标. 要求计算在时间间隔 $[a,b]$ 中克服重力所做的功.

功 $A(\alpha,\beta)$ 是区间 $[\alpha,\beta]\subset[a,b]$ 的可加函数.

当常力 $\boldsymbol{F}$ 作用于以常速度 $\boldsymbol{v}$ 运动的物体时, 在时间 h 内所做的功为 $\langle\boldsymbol{F},\boldsymbol{v}h\rangle=\langle\boldsymbol{F},\boldsymbol{v}\rangle h$. 因此, 可以想象, 成立以下估计

$$\inf_{t\in[\alpha,\beta]}\langle\boldsymbol{F}(p(t)),\boldsymbol{v}(t)\rangle(\beta-\alpha)\leqslant A(\alpha,\beta)\leqslant\sup_{t\in[\alpha,\beta]}\langle\boldsymbol{F}(p(t)),\boldsymbol{v}(t)\rangle(\beta-\alpha),$$

其中 $\boldsymbol{v}(t)$ 是物体在时刻 t 的速度, $p(t)$ 是它在时刻 t 在空间所处的位置, 而 $\boldsymbol{F}(p(t))$ 是在 $p=p(t)$ 点处作用于物体的力.

如果函数 $\langle \boldsymbol{F}(p(t)), \boldsymbol{v}(t)\rangle$ 可积, 则根据命题 1, 应当认为

$$A(a,b)=\int_a^b \langle \boldsymbol{F}(p(t)), \boldsymbol{v}(t)\rangle dt.$$

在上述具体情形, $\boldsymbol{v}(t)=(\dot{x}(t),\dot{y}(t),\dot{z}(t))$. 如果记 $\boldsymbol{r}(t)=(x(t),y(t),z(t))$, 则根据万有引力定律, 得

$$\boldsymbol{F}(p)=\boldsymbol{F}(x,y,z)=G\frac{mM}{|\boldsymbol{r}|^3}\boldsymbol{r}=\frac{GmM}{(x^2+y^2+z^2)^{3/2}}(x,y,z),$$

其中 M 是地球的质量, 并将坐标原点放在地球的中心.

这样一来,

$$\langle \boldsymbol{F}, \boldsymbol{v}\rangle(t)=GmM\frac{x(t)\dot{x}(t)+y(t)\dot{y}(t)+z(t)\dot{z}(t)}{(x^2(t)+y^2(t)+z^2(t))^{3/2}},$$

因此,

$$\begin{aligned}\int_a^b \langle \boldsymbol{F}, \boldsymbol{v}\rangle(t)dt &= \frac{1}{2}GmM\int_a^b \frac{(x^2(t)+y^2(t)+z^2(t))'}{(x^2(t)+y^2(t)+z^2(t))^{3/2}}dt \\ &= -\frac{GmM}{(x^2(t)+y^2(t)+z^2(t))^{1/2}}\bigg|_a^b = -\frac{GmM}{|\boldsymbol{r}(t)|}\bigg|_a^b.\end{aligned}$$

于是,

$$A(a,b)=\frac{GmM}{|\boldsymbol{r}(a)|}-\frac{GmM}{|\boldsymbol{r}(b)|}.$$

我们揭示了所求的功只依赖于 $|\boldsymbol{r}(a)|, |\boldsymbol{r}(b)|$ 这样一个事实, 即它只依赖于物体 m 在所考察的时间间隔 $[a,b]$ 的初始时刻和结束时刻离地球中心的距离.

令

$$U(r)=\frac{GM}{r}.$$

我们得到, 将质量为 m 的物体从半径为 r_0 的球面上任一点移动到半径为 r_1 的球面上任一点, 克服重力所做的功用公式

$$A_{r_0r_1}=m(U(r_0)-U(r_1))$$

计算.

函数 $U(r)$ 叫做牛顿位势. 如果以 R 表示地球半径, 那么, 由于 $\dfrac{GM}{R^2}=g$, 所以可以把函数 $U(r)$ 写成如下形式

$$U(r)=\frac{gR^2}{r}.$$

由此可以得出为脱离地球重力场, 准确地说, 为使质量为 m 的物体从地球表面飞到距地心无限远的地方, 所需要的功的表达式. 这个量显然是极限 $\lim\limits_{r\to+\infty} A_{Rr}$.

因此, 脱出功为

$$A = A_{R\infty} = \lim_{r\to+\infty} A_{Rr} = \lim_{r\to+\infty} m\left(\frac{gR^2}{R} - \frac{gR^2}{r}\right) = mgR.$$

练　习

1. 在图 51 上画出了沿横轴作用于处在该轴上 x 点处的试验粒子的力 $F = F(x)$ 的函数图像.
   a) 试在同一坐标系中画出这个力的势的草图.
   b) 画出力 $-F(x)$ 的势.
   c) 研究在 x_0 的各种可能的位置中怎样的位置是稳定平衡位置, 这与势的什么性质有关.

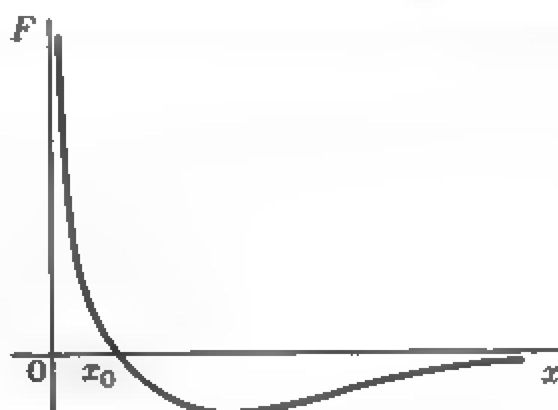


图 51

2. 在例 10 的结果的基础上计算物体飞离地球重力场所应具有的初速度 (关于地球的第二宇宙速度).
3. 在例 9 的基础上:
   a) 导出数学摆的振动方程 $R\ddot{\varphi} = g\sin\varphi$.
   b) 假定振幅很小, 试求它的近似解.
   c) 根据近似解确定摆的振动周期并与公式 (20) 的结果比较.
4. 有一个半径为 r 的轮子沿水平面以速度 v 均匀无滑动地滚动. 设在时刻 $t = 0$ 轮子最高点 A 的笛卡儿坐标是 $(0, 2r)$, 坐标系的横轴在指定的平面上, 其方向与速度 v 的方向相同.
   a) 写出点 A 的运动规律 $t \mapsto (x(t), y(t))$.
   b) 作为时间的函数, 求点 A 的速度.
   c) 画出点 A 的轨迹 (这条曲线叫*旋轮线*).
   d) 求该旋轮线一个拱的长度 (即该周期曲线一个周期的长度).
   e) 旋轮线有许多有趣的性质, 其中一个是惠更斯[①]发现的, 他说: 旋轮摆 (在旋轮槽中滚动的小球) 的振动周期不依赖于它从槽的最低点升起的高度. 试根据例 9 (参看讲述反常积分的下一节的问题 6) 做出证明.

① 惠更斯(Huygens)(1629—1695)—— 荷兰力学家、物理学家、数学家和天文学家.

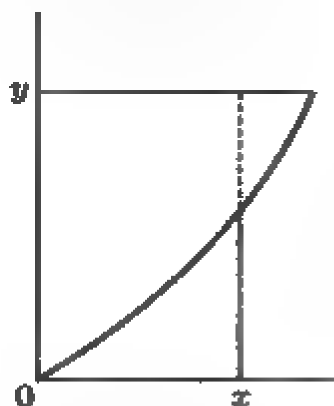


图 52

5. a) 见图 52. 设 $y=f(x)$ 和 $x=g(y)$ 是互逆的连续、非负函数, 且分别在 $x=0, y=0$ 时等于零. 试证成立不等式

$$xy \leqslant \int_0^x f(t)dt + \int_0^y g(t)dt.$$

b) 从 a) 导出杨格不等式

$$xy \leqslant \frac{1}{p}x^p + \frac{1}{q}y^q$$

其中 $x, y \geqslant 0, p, q > 0$ 且 $\frac{1}{p}+\frac{1}{q}=1$.

c) 在问题 a) 和 b) 的不等式中等号有什么几何意义?

6. 浦丰[①]问题. 数 π 可以用以下非正统的方法计算.

取一大张纸, 画满间隔为 h 的平行直线. 我们将一根长为 $l<h$ 的针完全随意的扔在纸上. 设扔了 N 次, 其中有 n 次扔下的针与一条直线相交. 如果 N 充分大, 则 $\pi \approx \frac{2l}{ph}$, 其中 $p=\frac{n}{N}$ 可以解释做扔下的针与一条直线相交的概率. 试从有关计算面积的几何知识出发给这个计算 π 的方法一个满意的解释.

§5. 反常积分

在上一节我们已经看到把黎曼积分概念作某些推广的必要性. 在那里通过对具体问题的分析, 我们形成了关于朝着什么方向推广黎曼积分概念和怎样推广的想法. 这一节我们来实现这些想法.

1. 反常积分的定义、例题和基本性质

定义 1 设函数 $x \mapsto f(x)$ 定义在区间 $[a,+\infty[$ 上, 而且在任何包含在这个区间内的闭区间 $[a,b]$ 上可积. 如果下式右端的极限存在, 则把量

$$\int_a^{+\infty} f(x)dx := \lim_{b\to+\infty}\int_a^b f(x)dx$$

[①]浦丰(Buffon)(1707—1788)—— 法国自然科学家.

叫做函数 $f(x)$ 在区间 $[a,+\infty[$ 上的反常黎曼积分, 或简单地叫做反常积分.

符号 $\int_a^{+\infty} f(x)dx$ 本身也叫做反常积分. 当上边的极限存在时就说它是收敛的反常积分. 反之, 则说它是发散的反常积分. 这样一来, 反常积分的收敛性问题就等价于它是否有定义的问题.

例 1 研究参数 α 取什么值时, 反常积分

$$\int_1^{+\infty} \frac{dx}{x^\alpha} \tag{1}$$

收敛或有定义.

因为

$$\int_1^b \frac{dx}{x^\alpha} = \begin{cases} \left.\dfrac{1}{1-\alpha}x^{1-\alpha}\right|_1^b, & \text{当}\alpha \neq 1\text{时}, \\ \left.\ln x\right|_1^b, & \text{当}\alpha = 1\text{时}, \end{cases}$$

所以, 仅当 $\alpha > 1$ 时, 极限 $\lim\limits_{b\to+\infty}\int_1^b \frac{dx}{x^\alpha} = \frac{1}{\alpha-1}$ 存在. 因此,

$$\int_1^{\infty} \frac{dx}{x^\alpha} = \frac{1}{\alpha-1}, \text{当}\alpha > 1.$$

对于 α 的其他值, 积分 (1) 发散, 也就是无定义.

定义 2 设函数 $x \mapsto f(x)$ 定义在区间 $[a,B[$ 上, 而在任何闭区间 $[a,b] \subset [a,B[$ 上可积. 如果下式右端的极限存在, 则把量

$$\int_a^B f(x)dx := \lim_{b\to B-0}\int_a^b f(x)dx$$

叫做函数 f 在区间 $[a,B[$ 上的反常积分.

这个定义的本质在于, 函数 f 可能在有限点 B 的任何邻域中都是无界的.

类似地, 如果函数 $x \mapsto f(x)$ 在区间 $]A,b]$ 上定义, 在任何闭区间 $[a,b] \subset]A,b]$ 上可积, 则可给出定义:

$$\int_A^b f(x)dx := \lim_{a\to A+0}\int_a^b f(x)dx.$$

同样地也可给出定义:

$$\int_{-\infty}^b f(x)dx := \lim_{a\to-\infty}\int_a^b f(x)dx.$$

例 2 研究参数 α 取什么值时, 积分

$$\int_0^1 \frac{dx}{x^\alpha} \tag{2}$$

收敛.

因为对于 $a \in]0,1]$, 有

$$\int_a^1 \frac{dx}{x^\alpha} = \begin{cases} \left.\dfrac{1}{1-\alpha} x^{1-\alpha}\right|_a^1, & \text{如果}\alpha \neq 1, \\ \ln x\Big|_a^1 & \text{, 如果}\alpha = 1. \end{cases}$$

所以仅当 $\alpha < 1$ 时, 极限

$$\lim_{a\to+0} \int_a^1 \frac{dx}{x^\alpha} = \frac{1}{1-\alpha}$$

存在.

这样, 积分 (2) 只对 $\alpha < 1$ 的情形有定义.

例 3

$$\int_{-\infty}^0 e^x dx = \lim_{a\to-\infty} \int_a^0 e^x dx = \lim_{a\to-\infty} \left(e^x\Big|_a^0\right) = \lim_{a\to-\infty} (1-e^a) = 1.$$

由于反常积分的收敛性问题, 无论说的是在无界区间上的反常积分, 还是在区间的一个端点附近无界的函数的反常积分, 解决的方法是一样的, 所以今后我们将把这两种情形放在一起讨论. 为此引进以下基本定义.

定义 3 设 $[a,\omega[$ 是有限或无限区间, $x \mapsto f(x)$ 是定义在这个区间上的函数, 而且它在每个闭区间 $[a,b] \subset [a,\omega[$ 上可积. 那么, 如果下式右端的极限存在, 则把

$$\int_a^\omega f(x)dx := \lim_{b\to\omega} \int_a^b f(x)dx \tag{3}$$

叫做函数 $f(x)$ 在区间 $[a,\omega[$ 上的反常积分.

如果没有相反的声明, 今后在研究反常积分 (3) 时, 总假定被积函数满足定义 3 的条件.

此外, 为确定起见, 暂时假定积分的反常性只与积分上限联系. 对反常性与下限联系的情形可完全同样地去研究.

从定义 3、积分的性质和极限的性质, 可以得到关于反常积分的以下结论.

命题 1 设 $x \mapsto f(x)$ 和 $x \mapsto g(x)$ 是定义在区间 $[a,\omega[$ 上且在任何闭区间 $[a,b] \subset [a,\omega[$ 上可积的函数. 设对它们能定义反常积分

$$\int_a^\omega f(x)dx, \tag{4}$$

$$\int_a^\omega g(x)dx. \tag{5}$$

那么,

a) 如果 $\omega \in \mathbb{R}, f \in \mathcal{R}[a,\omega]$, 则积分 (4) 的值, 无论理解做反常积分 还是常义积分都是一样的.

b) 对于任何 $\lambda_1, \lambda_2 \in \mathbb{R}$, 函数 $(\lambda_1 f + \lambda_2 g)(x)$ 在反常积分意义下在 $[a,\omega[$ 上可积, 而且成立等式

$$\int_a^\omega (\lambda_1 f + \lambda_2 g)(x)dx = \lambda_1 \int_a^\omega f(x)dx + \lambda_2 \int_a^\omega g(x)dx.$$

c) 如果 $c \in [a,\omega[$, 则

$$\int_a^\omega f(x)dx = \int_a^c f(x)dx + \int_c^\omega f(x)dx.$$

d) 如果 $\varphi : [\alpha,\gamma[\to [a,\omega[$ 是光滑、严格单调映射, 且 $\varphi(\alpha) = a$, 当 $\beta \in [\alpha,\gamma[$ 且 $\beta \to \gamma$ 时有 $\varphi(\beta) \to \omega$, 那么, 函数 $t \mapsto (f \circ \varphi)(t)\varphi'(t)$ 在 $[\alpha,\gamma[$ 上的反常积分存在且成立等式

$$\int_a^\omega f(x)dx = \int_\alpha^\gamma (f \circ \varphi)(t)\varphi'(t)dt.$$

◀a) 由函数

$$\mathcal{F}(b) = \int_a^b f(x)dx$$

在区间 $[a,\omega]$ 上的连续性 (因为 $f \in \mathcal{R}[a,\omega]$) 即得证.

b) 对于 $b \in [a,\omega[$, 有

$$\int_a^b (\lambda_1 f + \lambda_2 g)(x)dx = \lambda_1 \int_a^b f(x)dx + \lambda_2 \int_a^b g(x)dx.$$

由此即得证.

c) 对于任何 $b, c \in [a,\omega[$, 有

$$\int_a^b f(x)dx = \int_a^c f(x)dx + \int_c^b f(x)dx.$$

由此即得证.

d) 利用公式

$$\int_{a=\varphi(\alpha)}^{b=\varphi(\beta)} f(x)dx = \int_\alpha^\beta (f \circ \varphi)(t)\varphi'(t)dt,$$

即定积分的变量替换公式, 即得证. ▶

注 1 对命题 1 中所述反常积分的性质, 还应补充非常有用的反常积分的分部积分规则. 这条规则如下:

如果 $f, g \in C^{(1)}[a, \omega[$, 且存在极限 $\lim\limits_{\substack{x \to \omega \\ x \in [a,\omega[}} (f \cdot g)(x)$, 则函数 $f \cdot g'$ 和 $f' \cdot g$ 在区间 $[a, \omega[$ 上在反常积分意义下同时可积或不可积, 且当它们可积时成立等式

$$\int_a^\omega (f \cdot g')(x)dx = (f \cdot g)(x)\Big|_a^\omega - \int_a^\omega (f' \cdot g)(x)dx,$$

其中

$$(f \cdot g)(x)\Big|_a^\omega = \lim_{\substack{x \to \omega \\ x \in [a,\omega[}} (f \cdot g)(x) - (f \cdot g)(a).$$

◀ 这个结果可从常义积分的分部积分公式

$$\int_a^b (f \cdot g')(x)dx = (f \cdot g)\Big|_a^b - \int_a^b (f' \cdot g)(x)dx$$

得到. ▶

注 2 从命题 1 的 c) 看出, 反常积分

$$\int_a^\omega f(x)dx, \quad \int_c^\omega f(x)dx$$

同时收敛或发散. 这样一来, 在反常积分中跟在级数中一样, 收敛性与级数或积分的前段无关.

根据这条原理, 有时在提出反常积分的收敛性问题时, 干脆不写在其附近积分不具有奇性的那个积分限.

在这个约定下, 例 1,2 所得到的结果可以写成:

积分 $\displaystyle\int^{+\infty} \frac{dx}{x^\alpha}$ 只在 $\alpha > 1$ 时收敛;

积分 $\displaystyle\int_{+0} \frac{dx}{x^\alpha}$ 只在 $\alpha < 1$ 时收敛.

后边这个积分中的记号 $+0$ 表示积分区间是 $x > 0$.

利用变量替换, 从最后这个积分立刻推出

积分 $\displaystyle\int_{x_0+0} \frac{dx}{(x - x_0)^\alpha}$ 只在 $\alpha < 1$ 时收敛.

2. 反常积分收敛性的研究

a. 柯西判别法 根据定义 3, 反常积分 (3) 的收敛性等价于函数

$$\mathcal{F}(b) = \int_a^b f(x)dx \tag{6}$$

当 $b \to \omega, b \in [a, \omega[$ 时的极限的存在性.

因此, 下述断言成立.

命题 2 (反常积分收敛性的柯西判别法) 如果函数 $x \mapsto f(x)$ 定义在区间 $[a,\omega[$ 上, 而且在任何闭区间 $[a,b] \subset [a,\omega[$ 上可积, 则当且仅当对任何 $\varepsilon > 0$ 存在 $B \in [a,\omega[$ 使对一切 $b_1, b_2 \in [a,\omega[, B < b_1, B < b_2$, 成立关系

$$\left|\int_{b_1}^{b_2} f(x)dx\right| < \varepsilon$$

时, 积分 $\int_a^\omega f(x)dx$ 收敛.

◀ 事实上,

$$\int_{b_1}^{b_2} f(x)dx = \int_a^{b_2} f(x)dx - \int_a^{b_1} f(x)dx = \mathcal{F}(b_2) - \mathcal{F}(b_1).$$

因此, 命题中的条件正是函数 $\mathcal{F}(b)$ 当 $b \to \omega, b \in [a,\omega[$ 时极限存在性的柯西判别准则中的充分必要条件. ▶

b. 反常积分的绝对收敛性

定义 4 称反常积分 $\int_a^\omega f(x)dx$ 是绝对收敛的, 如果积分 $\int_a^\omega |f(x)|dx$ 收敛.

根据不等式

$$\left|\int_{b_1}^{b_2} f(x)dx\right| \leqslant \left|\int_{b_1}^{b_2} |f(x)|dx\right|$$

以及命题 2, 可以推出: 绝对收敛的反常积分必收敛.

绝对收敛性的研究归结为非负函数的积分的收敛性的研究. 而对于非负函数的情形我们有

命题 3 如果函数 f 满足定义 3 的条件且 $f(x) \geqslant 0$ 在 $[a,\omega[$ 上成立, 则当且仅当函数 (6) 在 $[a,\omega[$ 上有界时, 反常积分 (3) 存在.

◀ 事实上, 如果 $f(x) \geqslant 0$ 在 $[a,\omega[$ 上成立, 则函数 (6) 在 $[a,\omega[$ 上不减, 因此, 当且仅当这个函数有界时, 它当 $b \to \omega, b \in [a,\omega[$ 时有极限. ▶

作为应用这个命题的例子, 我们考察它的下述推论.

推论 (级数收敛性的积分准则) 如果 $x \mapsto f(x)$ 是定义在 $[1,+\infty[$ 上的非负、不增且在每个闭区间 $[1,b] \subset [1,+\infty[$ 上可积的函数, 则级数

$$\sum_{n=1}^{\infty} f(n) = f(1) + f(2) + \cdots$$

和积分

$$\int_1^{+\infty} f(x)dx$$

同时收敛或同时发散.

◀ 根据假设条件, 对于任意 $n \in \mathbb{N}$ 成立不等式

$$f(n+1) \leqslant \int_n^{n+1} f(x)dx \leqslant f(n).$$

对这个不等式求和, 我们得到

$$\sum_{n=1}^{k} f(n+1) \leqslant \int_1^{k+1} f(x)dx \leqslant \sum_{n=1}^{k} f(n)$$

或

$$s_{k+1} - f(1) \leqslant \mathcal{F}(k+1) \leqslant s_k,$$

其中 $s_k = \sum\limits_{n=1}^{k} f(n)$, 而 $\mathcal{F}(b) = \int_1^b f(x)dx$. 由于 s_k 和 $\mathcal{F}(b)$ 都是各自的自变量的不减函数, 所以由这些不等式即可导出我们的断言. ▶

特别地, 现在可以断定, 例 1 的结果等价于级数

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$$

只在 $\alpha > 1$ 时收敛.

命题 3 的最常用的推论是下述定理.

定理 (反常积分的比较定理) *设函数 $x \mapsto f(x), x \mapsto g(x)$ 在区间 $[a, \omega[$ 上定义, 且在任何闭区间 $[a, b] \subset [a, \omega[$ 上可积.*

如果在 $[a, \omega[$ 上有

$$0 \leqslant f(x) \leqslant g(x),$$

则从积分 (5) 的收敛性可以导出积分 (4) 的收敛性, 而且成立不等式

$$\int_a^\omega f(x)dx \leqslant \int_a^\omega g(x)dx;$$

而积分 (4) 的发散性可以导出积分 (5) 的发散性.

◀ 从定理的条件, 以及对于任何 $b \in [a, \omega[$ 常义黎曼积分的不等式, 我们有

$$\mathcal{F}(b) = \int_a^b f(x)dx \leqslant \int_a^b g(x)dx = \mathcal{G}(b).$$

由于 $\mathcal{F}, \mathcal{G}$ 都是 $[a, \omega[$ 上的不减函数, 所以从上面的不等式和命题 3 就推出欲证的结论. ▶

注 3 如果代替定理中的不等式 $0 \leqslant f(x) \leqslant g(x)$, 假定它们在 $[a,\omega[$ 上非负且当 $x \to \omega, x \in [a,\omega[$ 时是同阶的, 亦即存在正常数 c_1, c_2 使

$$c_1 f(x) \leqslant g(x) \leqslant c_2 f(x).$$

这时, 注意到反常积分的线性性质, 从上述定理推出: 积分 (4) 和 (5) 同时收敛或同时发散.

例 4 积分 $\int^{+\infty} \frac{\sqrt{x}dx}{\sqrt{1+x^4}}$ 收敛, 因为 $\frac{\sqrt{x}}{\sqrt{1+x^4}} \sim \frac{1}{x^{3/2}}$ 当 $x \to +\infty$.

例 5 积分 $\int_1^{+\infty} \frac{\cos x}{x^2} dx$ 绝对收敛, 因为 $\left|\frac{\cos x}{x^2}\right| \leqslant \frac{1}{x^2}$ 当 $x \geqslant 1$, 从而,

$$\int_1^{+\infty} \left|\frac{\cos x}{x^2}\right| dx \leqslant \int_1^{+\infty} \frac{dx}{x^2} = 1.$$

例 6 积分 $\int_1^{+\infty} e^{-x^2} dx$ 收敛, 因为 $0 < e^{-x^2} < e^{-x}$ 当 $x > 1$ 时成立, 从而

$$\int_1^{+\infty} e^{-x^2} dx < \int_1^{+\infty} e^{-x} dx = \frac{1}{e}.$$

例 7 积分 $\int^{+\infty} \frac{dx}{\ln x}$ 发散, 因为 $\frac{1}{\ln x} > \frac{1}{x}$ 当 x 充分大时成立.

例 8 欧拉积分 $\int_0^{\pi/2} \ln \sin x dx$ 收敛, 因为 $|\ln \sin x| \sim |\ln x| < 1/\sqrt{x}$ 当 $x \to +0$.

例 9 当 $0 \leqslant k^2 < 1$ 时, 椭圆积分

$$\int_0^1 \frac{dx}{\sqrt{(1-x^2)(1-k^2x^2)}}$$

收敛, 因为 $\sqrt{(1-x^2)(1-k^2x^2)} \sim \sqrt{2(1-k^2)}(1-x)^{1/2}$ 当 $x \to 1-0$.

例 10 积分

$$\int_0^{\varphi} \frac{d\theta}{\sqrt{\cos\theta - \cos\varphi}}$$

收敛, 因为

$$\sqrt{\cos\theta - \cos\varphi} = \sqrt{2\sin\frac{\varphi+\theta}{2}\sin\frac{\varphi-\theta}{2}} \sim \sqrt{\sin\varphi}(\varphi-\theta)^{1/2},$$

当 $\theta \to \varphi - 0$.

例 11 当 $0<\varphi_0<\pi$ 时, 积分

$$T=2\sqrt{\frac{L}{g}}\int_0^{\varphi_0}\frac{d\psi}{\sqrt{\sin^2\frac{\varphi_0}{2}-\sin^2\frac{\psi}{2}}} \tag{7}$$

收敛, 因为 $\psi\to\varphi_0-0$ 时, 有

$$\sqrt{\sin^2\frac{\varphi_0}{2}-\sin^2\frac{\psi}{2}}\sim\sqrt{\sin\varphi_0}(\varphi_0-\psi)^{1/2} \tag{8}$$

关系 (7) 是数学摆的振动周期对其长 L 和初始倾角的依赖关系. 所谓初始倾角即初始位置对穿过摆动轨迹最低点的半径的倾斜角. 公式 (7) 是上一节公式 (17) 的初等变形.

摆可以想象成, 譬如, 一根没有重量的杆, 它的一端是铰接固定的, 而另一端是自由的, 但系着一个质点.

这时, 关于任何初始角度 $\varphi_0\in[0,\pi]$ 都可进行讨论. 当 $\varphi_0=0$ 和 $\varphi_0=\pi$ 时, 摆根本不会振动. 在第一种情形, 摆处于稳定平衡状态; 而在第二种情形, 它处于不稳定平衡状态.

有趣的是, 从 (7) 和 (8) 不难得到, 当 $\varphi_0\to\pi-0$ 时, 有 $T\to\infty$, 也就是说, 摆的振动周期随着其初始位置无限接近其上方 (不稳定) 平衡位置而无限地增大.

c. 反常积分的条件收敛性

定义 5 反常积分如果收敛, 但不绝对收敛, 则说它是*条件收敛的反常积分*.

例 12 利用注 1, 根据反常积分的分部积分公式看出, 在

$$\int_{\pi/2}^{+\infty}\frac{\sin x}{x}dx=-\frac{\cos x}{x}\Big|_{\pi/2}^{+\infty}-\int_{\pi/2}^{+\infty}\frac{\cos x}{x^2}dx=-\int_{\pi/2}^{+\infty}\frac{\cos x}{x^2}dx$$

中, 如果最后的积分收敛, 则左边的积分也收敛. 但是, 由例 5 可知, $\int_{\pi/2}^{+\infty}\frac{\cos x}{x^2}dx$ 收敛, 从而积分

$$\int_{\pi/2}^{+\infty}\frac{\sin x}{x}dx \tag{9}$$

也收敛.

另外, 积分 (9) 不绝对收敛. 事实上, 对于 $b\in\left[\frac{\pi}{2},+\infty\right[$, 我们有

$$\int_{\pi/2}^{b}\left|\frac{\sin x}{x}\right|dx\geqslant\int_{\pi/2}^{b}\frac{\sin^2 x}{x}dx=\frac{1}{2}\int_{\pi/2}^{b}\frac{dx}{x}-\frac{1}{2}\int_{\pi/2}^{b}\frac{\cos 2x}{x}dx. \tag{10}$$

可以用分部积分法验证积分 $\int_{\pi/2}^{+\infty}\frac{\cos 2x}{x}dx$ 是收敛的. 因此, 当 $b\to+\infty$ 时, (10) 右端的差趋于 $+\infty$. 这样, 根据估计式 (10), 积分 (9) 不绝对收敛.

现在引进反常积分收敛性的一个特殊准则, 它基于第二中值定理, 就是说, 它本质上是以分部积分公式为基础的.

命题 4 (积分收敛性的阿贝尔–狄利克雷准则) 设 $x \mapsto f(x), x \mapsto g(x)$ 是定义在区间 $[a,\omega[$ 上并在任何闭区间 $[a,b] \subset [a,\omega[$ 上可积的函数. 设 g 是单调函数.

那么, 为使反常积分

$$\int_a^\omega (f \cdot g)(x)dx \tag{11}$$

收敛, 只需成立.

$\alpha_1)$ 积分 $\int_a^\omega f(x)dx$ 收敛,

$\beta_1)$ 函数 g 在 $[a,\omega[$ 上单调、有界或者成立,

$\alpha_2)$ 函数 $\mathcal{F}(b) = \int_a^b f(x)dx$ 在 $[a,\omega[$ 上有界,

$\beta_2)$ 函数 $g(x)$ 当 $x \to \omega, x \in [a,\omega[$ 时单调趋近于零.

◀ 对于任何 $b_1, b_2 \in [a,\omega[$, 根据第二中值定理, 有

$$\int_{b_1}^{b_2} (f \cdot g)(x)dx = g(b_1)\int_{b_1}^{\xi} f(x)dx + g(b_2)\int_{\xi}^{b_2} f(x)dx,$$

其中 ξ 是 b_1 和 b_2 之间的一个点. 由此, 根据柯西判别法 (命题 2) 得到: 如果上述两组条件之一成立, 积分 (11) 就收敛. ▶

3. 具有几个奇异点的反常积分 到现在为止, 我们只讨论了只有一个奇异点的反常积分, 它的被积函数在一个积分限附近无界, 或有一个积分限无界. 现在, 我们指出反常积分其他可能的类型.

如果两个积分限各是上述两种奇异类型中的一种, 则定义

$$\int_{\omega_1}^{\omega_2} f(x)dx := \int_{\omega_1}^{c} f(x)dx + \int_{c}^{\omega_2} f(x)dx, \tag{12}$$

其中 c 是区间 $]\omega_1,\omega_2[$ 中任一点. 我们假定 (12) 右边的每个反常积分都收敛. 如若不然, 则称 (12) 左边的积分发散.

根据注 3 和反常积分的可加性, 定义 (12) 在下述意义下是合理的: 它并不依赖于点 $c \in]\omega_1,\omega_2[$ 的选取.

例 13

$$\int_{-1}^{1} \frac{dx}{\sqrt{1-x^2}} = \int_{-1}^{0} \frac{dx}{\sqrt{1-x^2}} + \int_{0}^{1} \frac{dx}{\sqrt{1-x^2}}$$
$$= \arcsin x\Big|_{-1}^{0} + \arcsin x\Big|_{0}^{1} = \arcsin x\Big|_{-1}^{1} = \pi.$$

例 14 积分

$$\int_{-\infty}^{\infty} e^{-x^2} dx$$

叫欧拉-泊松积分有时也叫高斯积分. 显然, 它在上述意义下收敛. 稍后将证明它的值等于 $\sqrt{\pi}$.

例 15 积分

$$\int_0^{+\infty} \frac{dx}{x^\alpha}$$

发散, 因为对于任何 α, 在积分

$$\int_0^1 \frac{dx}{x^\alpha} \quad 和 \quad \int_1^{+\infty} \frac{dx}{x^\alpha}$$

中至少有一个是发散的.

例 16 积分

$$\int_0^{+\infty} \frac{\sin x}{x^\alpha} dx$$

收敛, 如果积分

$$\int_0^1 \frac{\sin x}{x^\alpha} dx \quad 和 \quad \int_1^{+\infty} \frac{\sin x}{x^\alpha} dx$$

都收敛.

当 $\alpha < 2$, 第一个积分收敛, 因为 $\frac{\sin x}{x^\alpha} \sim \frac{1}{x^{\alpha-1}}$ 当 $x \to +0$. 而当 $\alpha > 0$ 时, 第二个积分收敛. 这可用分部积分法, 类似在例 12 中做的那样去直接验证; 也可以引用阿贝尔-狄利克雷判别法. 这样一来, 原来的积分当 $0 < \alpha < 2$ 时有意义.

当被积函数在积分区间 $[a, b]$ 的一个内点 ω 的邻域中无界时, 定义

$$\int_a^b f(x)dx := \int_a^\omega f(x)dx + \int_\omega^b f(x)dx, \tag{13}$$

这里要求右边的两个积分都存在.

例 17 在 (13) 的意义下

$$\int_{-1}^1 \frac{dx}{\sqrt{|x|}} = 4.$$

例 18 积分 $\int_{-1}^1 \frac{dx}{x}$ 无定义.

关于在积分区间的内点 ω 的邻域中无界函数的积分, 还有与 (13) 不同的约定, 这就是

$$\text{V.P.} \int_a^b f(x)dx := \lim_{\delta \to +0} \left(\int_a^{\omega-\delta} f(x)dx + \int_{\omega+\delta}^b f(x)dx \right), \tag{14}$$

这里要求右边的极限存在. 沿用柯西的说法, 这个极限叫做主值意义下的积分.为了区别定义 (13) 和 (14), 在主值积分的情形, 积分号前边写上法文 Valeur Principal (主值) 这两个字开头的字母 V.P. 也使用英文记号 P.V. (源于 Principal Value)

在这个定义下, 我们有

例 19 $\text{V.P.}\int_{-1}^{1}\frac{dx}{x}=0.$

还有如下定义:

$$\text{V.P.}\int_{-\infty}^{+\infty}f(x)dx:=\lim_{R\to+\infty}\int_{-R}^{R}f(x)dx. \tag{15}$$

例 20 $\text{V.P.}\int_{-\infty}^{+\infty}xdx=0.$

最后, 如果在积分区间上有几个 (有限多个) 这种或那种奇异点, 它们位于区间内部或与区间端点重合, 则用这些奇异点把积分区间分成有限多个小区间, 其中每一个都只有一种奇异性的一个奇异点. 在原区间上的积分定义作分割后的诸小区间上的积分之和.

可以验证, 这样计算所得的结果不依赖于分割区间的方法.

例 21 现在可以把积分对数的精确定义写成以下形式:

$$\operatorname{li}x=\begin{cases}\displaystyle\int_0^x\frac{dt}{\ln t}, & \text{如果}0<x<1,\\ \displaystyle\text{V.P.}\int_0^x\frac{dt}{\ln t}, & \text{如果}1<x.\end{cases}$$

在后一种情形, 符号 V.P. 涉及的是区间 $]0,x]$ 上唯一的内部奇异点1. 我们注意到, 在定义 (13) 的意义下这个积分是不收敛的.

练　习

1. 试证: 函数

a) $\operatorname{Si}x=\int_0^x\frac{\sin t}{t}dt$(积分正弦) 是在 $\mathbb{R}$ 上定义的奇函数, 而且当 $x\to+\infty$ 时它有极限.

b) $\operatorname{si}x=-\int_x^{\infty}\frac{\sin t}{t}dt$ 在 $\mathbb{R}$ 上定义, 而且与 $\operatorname{Si}x$ 只差一个常数.

c) $\operatorname{Ci}x=-\int_x^{\infty}\frac{\cos t}{t}dt$(积分余弦) 对充分大的 x 可以用近似公式 $\operatorname{Ci}x\approx\frac{\sin x}{x}$ 计算. 试估计使绝对误差小于 10^{-4} 的那些 x 值的范围.

2. 试证:

a) 积分

$$\int_1^{+\infty}\frac{\sin x}{x^\alpha}dx,\int_1^{+\infty}\frac{\cos x}{x^\alpha}dx$$

只当 $\alpha > 0$ 时收敛, 而且只当 $\alpha > 1$ 时绝对收敛.

b) 菲涅耳积分

$$C(x) = \frac{1}{\sqrt{2}} \int_0^{\sqrt{x}} \cos t^2 dt,$$

$$S(x) = \frac{1}{\sqrt{2}} \int_0^{\sqrt{x}} \sin t^2 dt$$

在区间 $]0, +\infty[$ 上是无穷可微函数, 而且当 $x \to +\infty$ 时, 它们都有极限.

3. 试证:

a) 第一类椭圆积分

$$F(k, \varphi) = \int_0^{\sin \varphi} \frac{dt}{\sqrt{(1-t^2)(1-k^2t^2)}}$$

当 $0 \leqslant k < 1, 0 \leqslant \varphi \leqslant \pi/2$ 时有定义, 而且可以化成

$$F(k, \varphi) = \int_0^{\varphi} \frac{d\psi}{\sqrt{1 - k^2 \sin^2 \psi}}$$

的形式.

b) 第一类全椭圆积分

$$K(k) = \int_0^{\pi/2} \frac{d\psi}{\sqrt{1 - k^2 \sin^2 \psi}}$$

当 $k \to 1 - 0$ 时无限增加.

4. 试证:

a) 积分指数

$$\mathrm{Ei}(x) = \int_{-\infty}^{x} \frac{e^t}{t} dt$$

对于 $x < 0$ 有定义且无穷可微.

b) 当 $x \to +\infty$ 时, 有

$$-\mathrm{Ei}(x) = \frac{e^{-x}}{x} \left(1 - \frac{1}{x} + \frac{2!}{x^2} - \cdots + (-1)^n \frac{n!}{x^n} + o\left(\frac{1}{x^n}\right)\right).$$

c) 对于任何 $x \in \mathbb{R}$, 级数 $\sum\limits_{n=0}^{\infty} (-1)^n \frac{n!}{x^n}$ 都不收敛.

d) 当 $x \to +0$ 时, $\mathrm{li}\, x \sim \frac{x}{\ln x}$.(积分对数 $\mathrm{li}\, x$ 的定义可见于例 21)

5. 试证:

a) 函数

$$\mathrm{Fi}(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2} dt,$$

叫做概率误差积分, 常用符号 $\mathrm{erf}(x)$ 表示 (来源于英文 error function), 它是在 $\mathbb{R}$ 上定义的奇函数, 无穷可微且当 $x \to +\infty$ 时有极限.

b) 如果 a) 中所说的极限等于 1 (确实如此), 则当 $x \to +\infty$ 时, 有

$$\begin{aligned}\operatorname{erf}(x) &= \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt \\ &= 1-\frac{2}{\sqrt{\pi}}e^{-x^2}\left(\frac{1}{2x}-\frac{1}{2^2x^3}+\frac{1\cdot 3}{2^3x^5}-\frac{1\cdot 3\cdot 5}{2^4x^7}+o\left(\frac{1}{x^7}\right)\right).\end{aligned}$$

6. 试证:

a) 如果一个重的粒子在重力作用下沿着由参数形式 $x=x(\theta), y=y(\theta)$ 给出的曲线运动, 而且在时刻 $t=0$ 粒子的速度是 0, 位置在点 $x_0=x(\theta_0), y_0=y(\theta_0)$ 处, 那么, 曲线上参数为 θ 的点与粒子通过该点的时刻 t 有如下关系 (见 §4 的公式 (15)):

$$t=\pm\int_{\theta_0}^{\theta}\sqrt{\frac{(x'(\theta))^2+(y'(\theta))^2}{2g(y_0-y(\theta))}}d\theta,$$

如果 $y'(\theta_0)\neq 0$, 这个反常积分显然是收敛的 (在上述等式中符号的取法决定于 t 和 θ 是否具有同样的单调性, 而且, 如果 t 增长对应着 θ 增长, 显然应取正号).

b) 设槽的轮廓具有旋轮线形式:

$$\begin{aligned} x &= R(\theta+\pi+\sin\theta), \\ y &= -R(1+\cos\theta), \end{aligned} \quad |\theta|\leqslant\pi$$

则槽中运动粒子的振动周期不依赖于它开始滑动时的高度 $y_0=-R(1+\cos\theta_0)$, 永远等于 $4\pi\sqrt{\dfrac{R}{g}}$(见 §4 的问题 4).

第七章 多变量函数和它的极限与连续性

到目前为止, 我们研究的几乎只是数值函数 $x \mapsto f(x)$, 其中数 $f(x)$ 是由在函数定义域中给定的数 x 确定的.

但是很多使人感兴趣的量不是依赖于一个因素, 而是依赖于很多因素. 如果每个因素都可用数来描述, 即这些因素全体可用有序数组 $(x^1,\cdots,x^n)$ 表示, 其中每个数 $x^i(i=1,\cdots,n)$ 表示相应的因素. 设所研究的量本身也可用数表示, 记作 y. 那么, 它对确定它的因素的依赖关系可归结为用等式 $y=f(x^1,\cdots,x^n)$ 表示的有序数组 $(x^1,\cdots,x^n)$ 与数 y 的对应关系.

例如, 长方形面积是它的相邻两边长度的乘积. 一定量气体的体积用下述公式计算

$$V=R\frac{mT}{p},$$

其中 R—— 常数, m—— 质量, T—— 绝对温度, p —— 气体压力. 这样, 数值 V 依赖于由三个变数组成的有序数组 (m,T,p), 或者说, V 是三个变量 m,T,p 的函数.

我们已经学会了研究单变量函数. 我们现在的目标是能同样地学会研究多变量函数.

如同单变量函数一样, 多变量函数的研究也从描述它的定义域入手.

§1. 空间 $\mathbb{R}^m$和它的重要子集类

1. 集合 $\mathbb{R}^m$ 和 $\mathbb{R}^m$ 中的距离 我们约定用 $\mathbb{R}^m$ 表示由 m 个实数 $x^i\in\mathbb{R}(i=1,\cdots,m)$ 组成的所有有序数组 $(x^1,\cdots,x^m)$ 的集合.

每个这样的有序数组用一个字母 x 表示, 即 $x=(x^1,\cdots,x^m)$, 利用方便的几何术语, 称它为集合 $\mathbb{R}^m$ 中的*点*. 有序数组 $(x^1,\cdots,x^m)$ 中的数 x^i 称为点 $x=(x^1,\cdots,x^m)$ 的第 i 个*坐标*.

几何的类比还启示我们对集合 $\mathbb{R}^m$ 中的两点 $x_1=(x_1^1,\cdots,x_1^m), x_2=(x_2^1,\cdots,x_2^m)$ 按公式

$$d(x_1,x_2)=\sqrt{\sum_{i=1}^{m}(x_1^i-x_2^i)^2} \tag{1}$$

引进距离.

用公式 (1) 定义的函数

$$d:\mathbb{R}^m\times\mathbb{R}^m\longrightarrow\mathbb{R}$$

显然具有下面的性质:

a) $d(x_1,x_2)\geqslant 0$;

b) $(d(x_1,x_2)=0)\Leftrightarrow(x_1=x_2)$;

c) $d(x_1,x_2)=d(x_2,x_1)$;

d) $d(x_1,x_3)\leqslant d(x_1,x_2)+d(x_2,x_3)$.

最后这个不等式 (按几何说法仍然称为*三角形不等式*) 是闵可夫斯基不等式的特殊情形 (参看第五章 §4 第 2 段).

如果定义在集合 X 的元素对 (x_1,x_2) 上的函数具有性质 a),b),c),d), 则称它为 X 中的*度量或距离*.

集合 X 连同其中确定的距离一起称为*度量空间*.

这样一来, 我们赋予 $\mathbb{R}^m$ 以关系式 (1) 给出的度量以后, $\mathbb{R}^m$ 就成为度量空间了.

度量空间的一般知识, 读者可以在第九章 (第 II 卷) 中找到. 而在这里我们不想离开我们现在所必需的具体度量空间 $\mathbb{R}^m$.

在这一章中, 我们研究的唯一对象是带有距离 (1) 的集合 $\mathbb{R}^m$, 因此, 一般的度量空间定义暂时还不需要. 引入一般的度量空间仅仅是为了解释为什么对集合 $\mathbb{R}^m$ 和函数 (1) 分别使用了“空间”与“度量”这两个术语.

从关系式 (1) 得到, 当 $i\in\{1,\cdots,m\}$, 有

$$|x_1^i-x_2^i|\leqslant d(x_1,x_2)\leqslant\sqrt{m}\cdot\max_{1\leqslant i\leqslant m}|x_1^i-x_2^i|, \tag{2}$$

也就是说, 两点 $x_1,x_2\in\mathbb{R}^m$ 之间的距离很小当且仅当对应的坐标之差很小.

从式 (2) 或式 (1) 容易看到, 当 $m=1$ 时, 集合 $\mathbb{R}^1$ 完全与实数集是一样的, 它的点之间的距离也是以常规方式利用两数之差的模来计量的.

2. $\mathbb{R}^m$ 中的开集与闭集

定义 1 设 $\delta > 0$, 称集合

$$B(a;\delta) := \{x \in \mathbb{R}^m | d(a,x) < \delta\}$$

为以 a 为中心, 半径为 δ 的球或点 $a \in \mathbb{R}^m$ 的 δ 邻域.

定义 2 集合 $G \subset \mathbb{R}^m$ 称为 $\mathbb{R}^m$ 中的开集, 如果对于任意的点 $x \in G$, 存在一个球 $B(x;\delta)$, 使得 $B(x;\delta) \subset G$.

例 1 $\mathbb{R}^m$ 是 $\mathbb{R}^m$ 中的开集.

例 2 空集 $\varnothing$ 不含任何点, 因此, 可认为它满足定义 2, 即空集 $\varnothing$ 是 $\mathbb{R}^m$ 中的开集.

例 3 球 $B(a;r)$ 是 $\mathbb{R}^m$ 中的开集, 因此, 称它为开球.

事实上, 如果 $x \in B(a;r)$, 即 $d(a,x) < r$, 那么, 当 $0 < \delta < r - d(a,x)$ 时, 有 $B(x;\delta) \subset B(a;r)$, 这是因为

$$(\xi \in B(x;\delta)) \Longrightarrow (d(x,\xi) < \delta)$$
$$\Longrightarrow (d(a,\xi) < d(a,x) + d(x,\xi) < d(a,x) + r - d(a,x) = r).$$

例 4 集合 $G = \{x \in \mathbb{R}^m | d(a,x) > r\}$, 也就是与给定点 $a \in \mathbb{R}^m$ 的距离大于 r 的所有点构成的集合是 $\mathbb{R}^m$ 中的开集, 这可以同例 3 一样, 利用距离的三角形不等式加以证明.

定义 3 集合 $F \subset \mathbb{R}^m$ 称为 $\mathbb{R}^m$ 中的闭集, 如果它在 $\mathbb{R}^m$ 中的余集 $G = \mathbb{R}^m \backslash F$ 是 $\mathbb{R}^m$ 中的开集.

例 5 集合 $\bar{B}(a;r) := \{x \in \mathbb{R}^m | d(a,x) \leqslant r\}, r \geqslant 0$, 也就是与给定点 $a \in \mathbb{R}^m$ 的距离不大于 r 的所有点构成的集合, 是 $\mathbb{R}^m$ 中的闭集. 这可以从定义 3 与例 4 加以证明. 集合 $\bar{B}(a;r)$ 称为以 a 为中心、半径为 r 的闭球.

命题 1 a) 设 $\{G_\alpha, \alpha \in A\}$ 是 $\mathbb{R}^m$ 中任意多个开集构成的集合族, 则它的并集 $\bigcup_{\alpha \in A} G_\alpha$ 是 $\mathbb{R}^m$ 中的开集.

b) $\mathbb{R}^m$ 中有限多个开集的交集 $\bigcap_{i=1}^n G_i$ 是 $\mathbb{R}^m$ 中的开集.

a′) 设 $\{F_\alpha, \alpha \in A\}$ 是 $\mathbb{R}^m$ 中任意多个闭集构成的集合族, 则它的交集 $\bigcap_{\alpha \in A} F_\alpha$ 是 $\mathbb{R}^m$ 中的闭集.

b′) $\mathbb{R}^m$ 中有限多个闭集的并集 $\bigcup_{i=1}^n F_i$ 是 $\mathbb{R}^m$ 中的闭集.

◀ a) 若 $x \in \bigcup_{\alpha \in A} G_\alpha$, 则存在 $\alpha_0 \in A$, 使得 $x \in G_{\alpha_0}$, 因而存在点 x 的一个 δ 邻域 $B(x,\delta)$, 使得 $B(x;\delta) \subset G_{\alpha_0}$, 于是, $B(x;\delta) \subset \bigcup_{\alpha \in A} G_\alpha$.

b) 设 $x \in \bigcap_{i=1}^{n} G_i$, 则 $x \in G_i (i = 1, \cdots, n)$.

设 $\delta_1, \cdots, \delta_n$ 是这样的正数, 使得 $B(x; \delta_i) \subset G_i (i = 1, \cdots, n)$. 令 $\delta = \min \{\delta_1, \cdots, \delta_n\}$, 显然有 $\delta > 0$ 且 $B(x; \delta) \subset \bigcap_{i=1}^{n} G_i$.

a′) 我们要证明, 集合 $\bigcap_{\alpha \in A} F_\alpha$ 在 $\mathbb{R}^m$ 中的余集 $C\left(\bigcap_{\alpha \in A} F_\alpha\right)$ 是 $\mathbb{R}^m$ 中的开集.

事实上,

$$C\left(\bigcap_{\alpha \in A} F_\alpha\right) = \bigcup_{\alpha \in A} (CF_\alpha) = \bigcup_{\alpha \in A} G_\alpha.$$

其中 $G_\alpha = CF_\alpha$ 是 $\mathbb{R}^m$ 中的开集, 于是由 a) 证得 a′).

b′) 类似地, 从 b) 得到,

$$C\left(\bigcup_{i=1}^{n} F_i\right) = \bigcap_{i=1}^{n} (CF_i) = \bigcap_{i=1}^{n} G_i. \blacktriangleright$$

例 6 集合 $S(a; r) := \{x \in \mathbb{R}^m | d(a, x) = r\}, r \geqslant 0$, 叫做以 $a \in \mathbb{R}^m$ 为中心、半径为 r 的*球面*. $S(a; r)$ 在 $\mathbb{R}^m$ 中的余集由例 3 与例 4 知道是开集的并集, 再由命题 1 的 a) 知道它是开集, 因此, 球面 $S(a; r)$ 是 $\mathbb{R}^m$ 中的闭集.

定义 4 $\mathbb{R}^m$ 中包含给定点的开集, 称为这个点在 $\mathbb{R}^m$ 中的*邻域*.

特别地, 从例 3 得到, 某个点的 δ 邻域是这个点在 $\mathbb{R}^m$ 中的邻域.

定义 5 设点 $x \in \mathbb{R}^m$, 集合 $E \subset \mathbb{R}^m$.

如果点 x 连同它的某个邻域都含在 E 中, 则称点 x 是 E 的*内点*;
如果点 x 是 E 在 $\mathbb{R}^m$ 中的余集的内点, 则称点 x 是 E 的*外点*;
如果点 x 既不是 E 的内点, 也不是 E 的外点, 则称点 x 是 E 的*边界点*.

从这个定义知道, 集合的边界点的特征性质是: 它的任一邻域中既含有集合的点, 也含有不属于此集合的点.

例 7 球面$S(a; r), r > 0$, 既是开球$B(a; r)$ 的, 也是闭球$\bar{B}(a; r)$ 的边界点的集合.

例 8 点 $a \in \mathbb{R}^m$ 是集合 $\mathbb{R}^m \backslash a$ 的边界点, 这个集合没有外点.

例 9 球面 $S(a; r)$ 上的所有点都是这个球面的边界点. $S(a; r)$ 作为 $\mathbb{R}^m$ 的子集没有内点.

定义 6 点 $a \in \mathbb{R}^m$ 称为集合 $E \subset \mathbb{R}^m$ 的*极限点*, 如果点 a 的任一邻域 $O(a)$ 与 E 的交集 $E \cap O(a)$ 都是无限集.

定义 7 集合 $E \subset \mathbb{R}^m$ 和它的所有极限点组成的集的并集, 称为 E 在 $\mathbb{R}^m$ 中的*闭包*. 集合 E 的闭包常记作 $\bar{E}$.

例 10 $\bar{B}(a;r) = B(a;r) \cup S(a;r)$ 是开球$B(a;r)$ 的所有极限点的集合，因此 $\bar{B}(a;r)$ 不同于 $B(a;r)$，称为闭球.

例 11 $\bar{S}(a;r) = S(a;r)$.

我们不去说明上式成立的理由，而来证明以下有益的命题.

命题 2 *(F 是 $\mathbb{R}^m$ 中的闭集)$\Longleftrightarrow$(在 $\mathbb{R}^m$ 中有 $\bar{F} = F$).*

换句话说，F 是 $\mathbb{R}^m$ 中的闭集当且仅当它含有它所有的极限点.

◀ 设 F 是 $\mathbb{R}^m$ 中的闭集，$x \in \mathbb{R}^m, x \notin F$. 这时，开集 $G = \mathbb{R}^m \setminus F$ 是点 x 的邻域，它不含 F 的点，这就证明了，如果 $x \notin F$，则 x 不是 F 的极限点.

设 $F = \bar{F}$. 要证 $G = \mathbb{R}^m \setminus \bar{F}$ 是 $\mathbb{R}^m$ 中的开集. 如果 $x \in G$，则 $x \notin \bar{F}$，因此 x 不是 F 的极限点，这意味着存在点 x 的一个邻域，它仅含 F 的有限个点 $x_1, \cdots, x_n$. 因为 $x \notin F$，所以可构造点 x 的球形邻域 $O_1(x), \cdots, O_n(x)$，使得 $x_i \notin O_i(x)$，这时 $O(x) = \bigcap_{i=1}^{n} O_i(x)$ 是点 x 的邻域且不含 F 的点，即 $O(x) \subset \mathbb{R}^m \setminus F$，这就是说，$\mathbb{R}^m \setminus F = \mathbb{R}^m \setminus \bar{F} = G$ 是开集，因而 F 是 $\mathbb{R}^m$ 中的闭集. ▶

3. $\mathbb{R}^m$ 中的紧集

定义 8 集合 $K \subset \mathbb{R}^m$ 称为*紧集*，如果从任意的由 $\mathbb{R}^m$ 的开集构成的 K 的覆盖中总能选出有限覆盖.

例 12 闭区间 $[a,b] \subset \mathbb{R}^1$ 是 $\mathbb{R}^1$ 中的紧集，这可以由有限覆盖引理证明.

例 13 区间在 $\mathbb{R}^m$ 中的推广指的是集合

$$I = \{x \in \mathbb{R}^m | a^i \leqslant x^i \leqslant b^i, i = 1, \cdots, m\},$$

它也叫做m *维区间*，m 维长方体或 m 维平行多面体.

我们证明:I 是 $\mathbb{R}^m$ 中的紧集.

◀ 假定有 I 的某个开覆盖不含有限覆盖. 把 I 的每个坐标区间

$$I^i = \{x^i \in \mathbb{R} | a^i \leqslant x^i \leqslant b^i\} \quad (i = 1, \cdots, m)$$

二等分，就把 I 分成了 2^m 个 m 维区间，这些 m 维区间中至少有一个不具有有限覆盖，对这个 m 维区间仍采取上面的做法，又得到一个不具有有限覆盖的 m 维区间，继续这个过程，我们得到一个闭 m 维区间序列

$$I = I_1 \supset I_2 \supset \cdots \supset I_n \supset \cdots,$$

这个序列中任何一个闭 m 维区间都不具有有限覆盖. 如果记

$$I_n = \{x \in \mathbb{R}^m | a_n^i \leqslant x^i \leqslant b_n^i, i = 1, \cdots, m\},$$

那么, 按照上面的做法, 对于每个 $i \in \{1, \cdots, m\}$, 坐标区间 $a_n^i \leqslant x^i \leqslant b_n^i \quad (n = 1, 2, \cdots)$ 构成一个闭区间套, 它的长趋于零, 因此对每个 $i \in \{1, \cdots, m\}$, 存在一个属于所有这些区间的公共点 $\xi^i \in [a_n^i, b_n^i]$, 由此得到属于所有 m 维区间 $I = I_1, I_2, \cdots, I_n, \cdots$ 的点 $\xi = (\xi^1, \cdots, \xi^m)$. 因为 $\xi \in I$, 所以在覆盖 I 的开集族中, 存在一个开集 G, 使得 $\xi \in G$, 这就是说, 存在 $\delta > 0$, 使得 $B(\xi; \delta) \subset G$. 但是, 按照上面的做法及关系式 (2), 可以找到足码 N, 当 $n > N$ 时, 有 $I_n \subset B(\xi; \delta) \subset G$ 而这与 m 维区间 I_n 在给定覆盖族中没有有限覆盖是矛盾的. ►

命题 3 *如果 K 是 $\mathbb{R}^m$ 中的紧集, 则*

a) K 是 $\mathbb{R}^m$ 中的闭集;

b) K 中任何一个闭集都是紧集.

◄ a) 我们要证:K 的任一极限点 $a \in \mathbb{R}^m$ 都属于 K. 假设 $a \notin K$. 对于每个 $x \in K$, 可以构造这样的邻域 $G(x)$, 使得点 a 也存在一个邻域与 $G(x)$ 没有公共点. 所有这些邻域组成的集合族 $\{G(x)\}(x \in K)$, 是紧集 K 的一个开覆盖, 从中可选出有限覆盖 $G(x_1), \cdots, G(x_n)$. 如果 $O_i(a)$ 是点 a 的邻域, 且使 $G(x_i) \cap O_i(a) = \varnothing$, 那么, 集合 $O(a) = \bigcap_{i=1}^{n} O_i(a)$ 同样是点 a 的邻域, 而且显然有 $K \cap O(a) = \varnothing$. 这样一来, 点 a 不可能是 K 的极限点.

b) 设 F 是 $\mathbb{R}^m$ 中的闭集, 且 $F \subset K$. 设 $\{G_\alpha\}, \alpha \in A$ 是覆盖 F 的开集族, 再加上一个开集 $G = \mathbb{R}^m \backslash F$, 就构成 $\mathbb{R}^m$ 的开覆盖, 特别是 K 的开覆盖, 从中选出 K 的一个有限覆盖, 它同样覆盖了 F, 注意到 $G \cap F = \varnothing$, 因此, 由 G 属于这个有限覆盖, 可知, 去掉它以后, 我们仍然能从开集族 $\{G_\alpha\}(\alpha \in A)$ 中找到 F 的一个有限覆盖. ►

定义 9 量

$$d(E) := \sup_{x_1, x_2 \in E} d(x_1, x_2)$$

称为集合 $E \subset \mathbb{R}^m$ 的*直径*.

定义 10 集合 $E \subset \mathbb{R}^m$ 称为*有界集*, 如果它的直径是有限数.

命题 4 *如果 K 是 $\mathbb{R}^m$ 中的紧集, 则 K 是 $\mathbb{R}^m$ 中的有界集.*

◄ 任取 $a \in \mathbb{R}^m$, 考察一系列的球 $\{B(a; n)\} \quad (n = 1, 2, \cdots)$, 它们构成 $\mathbb{R}^m$ 的一个开覆盖, 因而也是 K 的一个开覆盖, 假如 K 不是有界集, 那么, 从这个开覆盖中不可能找出 K 的一个有限覆盖. ►

命题 5 *集合 $K \subset \mathbb{R}^m$ 是紧集当且仅当 K 是 $\mathbb{R}^m$ 中的有界闭集.*

◄ 这个命题的必要性我们已经在命题 3 与命题 4 中证明过了.

现在证明命题的充分性. 因为 K 是有界集, 所以, 可以找到一个 m 维区间 I 包含 K, 在例 13 中我们证明了 I 是 $\mathbb{R}^m$ 中的紧集, 又因 K 是包含在 I 中的闭集, 根

据命题 3 的 b) 知道 K 本身也是 $\mathbb{R}^m$ 中的紧集. ▶

练 习

1. 设 $E_1, E_2 \subset \mathbb{R}^m$, 量

$$d(E_1, E_2) := \inf_{x_1 \in E_1, x_2 \in E_2} d(x_1, x_2)$$

称为集合 E_1 与 E_2 之间的距离.

举例说明, 在 $\mathbb{R}^m$ 中存在没有公共点的闭集 E_1, E_2 使 $d(E_1, E_2) = 0$.

2. 证明:

a) 任一集合 $E \subset \mathbb{R}^m$ 的闭包 $\bar{E}$ 是 $\mathbb{R}^m$ 中的闭集.

b) 任一集合 $E \subset \mathbb{R}^m$ 的边界点的集合 ∂E 是 $\mathbb{R}^m$ 中的闭集.

c) 如果 G 是 $\mathbb{R}^m$ 中的开集, F 是 $\mathbb{R}^m$ 中的闭集, 则 $G \setminus F$ 是 $\mathbb{R}^m$ 中的开集.

3. 证明: 如果 $K_1 \supset K_2 \supset \cdots \supset K_n \supset \cdots$ 是一个紧集列套, 则 $\bigcap_{i=1}^{\infty} K_i \neq \varnothing$.

4. a) 在空间 $\mathbb{R}^k$ 中, 二维球面 S^2 和圆周 S^1 能不能有那样的位置关系: 从球面上任意点到圆周上任意点的距离都是一样的.

b) 对任意维球面 $S^m, S^n \subset \mathbb{R}^k$ 研究问题 a). 试问, 当 m, n, k 有什么关系时, 所说的位置关系成立.

§2. 多变量函数的极限与连续性

1. 函数的极限 在第三章中我们已经详细研究了定义在集合 X 上的实值函数 $f: X \to \mathbb{R}$ 关于 X 中的基 $\mathcal{B}$ 的极限运算.

在本节中, 我们要考察定义在集合 X 上而值域在 $\mathbb{R}^n (n \in \mathbb{N})$ 中的函数 $f: X \to \mathbb{R}^n$ 的极限理论.

我们从一般的基本定义开始.

定义 1 点 $A \in \mathbb{R}^n$ 称为映射$f: X \to \mathbb{R}^n$关于 X 中基 $\mathcal{B}$ 的极限, 如果对于点 A 的任一邻域 $V(A)$, 存在 $B \in \mathcal{B}$, 使得 $f(B) \subset V(A)$.

简言之,

$$\lim_{\mathcal{B}} f(x) = A := (\forall V(A) \exists B \in \mathcal{B} (f(B) \subset V(A))).$$

我们看到, 函数 $f: X \to \mathbb{R}^n$ 的极限定义与函数 $f: X \to \mathbb{R}$ 的极限定义是完全一致的, 如果对于任何 $n \in \mathbb{N}$ 我们都想像得出点 $A \in \mathbb{R}^n$ 的邻域是什么的话.

定义 2 映射$f: X \to \mathbb{R}^n$ 叫做有界映射, 如果集合 $f(X) \subset \mathbb{R}^n$ 在 $\mathbb{R}^n$ 中有界.

定义 3 设 $\mathcal{B}$ 是集合 X 中的基. 映射$f: X \to \mathbb{R}^n$ 叫做对基 $\mathcal{B}$ 最终有界映射, 如果存在 $B \in \mathcal{B}$, 使得 f 在 B 上有界.

由这些定义出发，容易证明我们在第三章中已经证明了的那些论断，即

函数 $f: X \to \mathbb{R}^n$ 关于 X 中的基 $\mathcal{B}$ 至多有一个极限；

若函数 $f: X \to \mathbb{R}^n$ 关于 X 中的基 $\mathcal{B}$ 有极限，则 f 对于基 $\mathcal{B}$ 是最终有界的.

直接了当地利用 $\mathbb{R}^n$ 中的度量，可以把定义 1 写成另外的形式，即

定义1′

$$(\lim_{\mathcal{B}} f(x) = A \in \mathbb{R}^n) := (\forall \varepsilon > 0 \exists B \in \mathcal{B} \forall x \in B (d(f(x), A) < \varepsilon))$$

或

定义1″

$$(\lim_{\mathcal{B}} f(x) = A \in \mathbb{R}^n) := (\lim_{\mathcal{B}} d(f(x), A) = 0).$$

映射 $f: X \to \mathbb{R}^n$ 的特征是：由于点 $y \in \mathbb{R}^n$ 是由 n 个实数组成的有序数组 $(y^1, \cdots, y^n)$，因此，函数 $f: X \to \mathbb{R}^n$ 等价于 n 个实值函数 $f^i: X \to \mathbb{R} (i = 1, \cdots, n)$，其中 $f^i(x) = y^i (i = 1, \cdots, n)$.

如果 $A = (A^1, \cdots, A^n), y = (y^1, \cdots, y^n)$，那么，成立不等式

$$|y^i - A^i| \leqslant d(y, A) \leqslant \sqrt{n} \cdot \max_{1 \leqslant i \leqslant n} |y^i - A^i|, \tag{1}$$

从它看出，

$$\lim_{\mathcal{B}} f(x) = A \Leftrightarrow \lim_{\mathcal{B}} f^i(x) = A^i \quad (i = 1, \cdots, n), \tag{2}$$

也就是说，$\mathbb{R}^n$ 中的收敛是按坐标收敛.

现在设 $X = \mathbb{N}$ —— 自然数集，而 $\mathcal{B}$ 是它的基 $k \to \infty, k \in \mathbb{N}$. 这时，函数 $f: \mathbb{N} \to \mathbb{R}^n$ 是空间 $\mathbb{R}^n$ 中的点列 $\{y_k\}, k \in \mathbb{N}$.

定义 4 点列 $\{y_k\}, k \in \mathbb{N}, y_k \in \mathbb{R}^n$ 称为*基本点列*，如果对于任意的 $\varepsilon > 0$，存在 $N \in \mathbb{N}$，使得对任意的 $k_1, k_2 > N$，有 $d(y_{k_1}, y_{k_2}) < \varepsilon$.

从不等式 (1) 看出，点列 $y_k = (y_k^1, \cdots, y_k^n) \in \mathbb{R}^n$ 是基本的当且仅当它的每一个坐标列 $\{y_k^i\}, k \in \mathbb{N}, i = 1, \cdots, n$ 是基本的.

由不等式 (1) 及数列收敛的柯西准则可以断言，$\mathbb{R}^n$ 中的点列是收敛的当且仅当它是基本的.

换句话说，柯西准则在 $\mathbb{R}^n$ 中是正确的.

今后，一个度量空间，如果其中每个基本列都有极限，则称它为*完备度量空间*. 这样，我们现在证明了对任意的 $n \in \mathbb{N}, \mathbb{R}^n$ 是完备度量空间.

定义 5 设 $E \subset X$，量

$$\omega(f; E) := d(f(E))$$

称为函数 $f: X \to \mathbb{R}^n$ 在 E 上的振幅, 其中 $d(f(E))$ 是集合 $f(E) \subset \mathbb{R}^n$ 的直径.

由此可见, 定义 5 是实值函数在集合 E 上的振幅概念的直接推广.

与 $\mathbb{R}^n$ 的完备性有关, 我们有下面关于函数 $f: X \to \mathbb{R}^n$ 极限存在性的柯西准则.

定理 1 *设 X 是一个集合, $\mathcal{B}$ 是 X 中的基, 函数 $f: X \to \mathbb{R}^n$ 关于基 $\mathcal{B}$ 有极限当且仅当对于任意的 $\varepsilon > 0$, 存在 $B \in \mathcal{B}$, 使得 f 在 B 上的振幅小于 ε.*

也就是

$$\exists \lim_{\mathcal{B}} f(x) \Leftrightarrow \forall \varepsilon > 0\ \exists B \in \mathcal{B}\ (\omega(f; B) < \varepsilon).$$

定理 1 的证明可以逐字逐句重复函数极限存在的柯西准则的证明 (第三章 §2 定理 4), 唯一需要改变的是把 $|f(x_1) - f(x_2)|$ 换为 $d(f(x_1), f(x_2))$.

定理 1 的正确性还可用另外方式加以证明, 这就是承认实值函数的柯西准则并利用关系式 (2) 和 (1).

对于函数值在 $\mathbb{R}^n$ 中的函数, 关于复合函数极限的重要定理也成立.

定理 2 *设 Y 是集合, $\mathcal{B}_Y$ 是 Y 中的基, 映射 $g: Y \to \mathbb{R}^n$ 关于基 $\mathcal{B}_Y$ 有极限.*

设 X 是集合, $\mathcal{B}_X$ 是 X 中的基, $f: X \to Y$ 是从 X 到 Y 的映射, 而且, 对于任意的 $B_Y \in \mathcal{B}_Y$, 存在 $B_X \in \mathcal{B}_X$, 使得 $f(B_X) \subset B_Y$.

在这些条件下, 映射 f 与 g 的复合映射 $g \circ f: X \to \mathbb{R}^n$ 关于基 $\mathcal{B}_X$ 有极限, 且

$$\lim_{\mathcal{B}_X} (g \circ f)(x) = \lim_{\mathcal{B}_Y} g(y).$$

为了证明定理 2, 或者重复第三章 §2 定理 5 的证明, 只要把 $\mathbb{R}$ 换成 $\mathbb{R}^n$, 或者利用第三章 §2 定理 5 及关系式 (2).

直到现在, 我们考察的函数 $f: X \to \mathbb{R}^n$ 的定义域 X 是一个抽象的集合, 今后, 首先使我们感兴趣的是 X 为空间 $\mathbb{R}^m$ 中的子集的情形.

像以前一样, 我们约定,

$U(a)$ —— 点 $a \in \mathbb{R}^m$ 的邻域;

$\mathring{U}(a)$ —— 点 $a \in \mathbb{R}^m$ 的去心邻域, 即 $\mathring{U}(a) := U(a) \backslash a$;

$U_E(a)$ —— 点 a 在集合 E 中的邻域, 即 $U_E(a) := E \cap U(a)$;

$\mathring{U}_E(a)$ —— 点 a 在集合 E 中的去心邻域, 即 $\mathring{U}_E(a) := E \cap \mathring{U}(a)$;

$x \to a$ —— 点 a 的去心邻域组成 $\mathbb{R}^m$ 中的基;

$x \to \infty$ —— 无穷远点邻域组成的基, 即由集合 $\mathbb{R}^m \backslash B(a; r)$ 组成的基;

$x \to a, x \in E$ 或 $(E \ni x \to a)$ —— 点 a 在集合 E 中的去心邻域组成的基, 且 a 是 E 的极限点;

$x \to \infty, x \in E$ 或 $(E \ni x \to \infty)$ —— 无穷远点在集合 E 中的邻域, 即 $E \backslash B(a; r)$ 组成的基, 且 E 是无界集.

如果 $f: E \to \mathbb{R}^n$ 是从集合 $E \subset \mathbb{R}^m$ 到 $\mathbb{R}^n$ 中的映射, 那么, 利用上面的记号可以使定义 1 表示的函数极限具体化:

$$\left(\lim_{E \ni x \to a} f(x) = A\right) := (\forall \varepsilon > 0\ \exists \overset{\circ}{U}_E(a)\ \forall x \in \overset{\circ}{U}_E(a)\ (d(f(x), A) < \varepsilon)).$$

上式也可写成另外的样子:

$$\left(\lim_{E \ni x \to a} f(x) = A\right) := (\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in E\ (0 < d(x, a) < \delta \Rightarrow d(f(x), A) < \varepsilon)).$$

其中 $d(x, a)$ 与 $d(f(x), A)$ 分别是空间 $\mathbb{R}^m$ 与 $\mathbb{R}^n$ 中所指的点之间的距离.

最后,

$$\left(\lim_{x \to \infty} f(x) = A\right) := (\forall \varepsilon > 0\ \exists B(a; r)\ \forall x \in \mathbb{R}^m \backslash B(a; r)\ (d(f(x), A) < \varepsilon)).$$

我们还约定, 对于映射 $f: X \to \mathbb{R}^n$, 术语 "对于基 $\mathcal{B}, f(x) \to \infty$" 指的是: 对于任意的球 $B(A; r) \subset \mathbb{R}^n$, 存在 $B \in \mathcal{B}$, 使得 $f(B) \subset \mathbb{R}^n \backslash B(A; r)$.

例 1 设 $x \mapsto \pi^i(x)$, 其中 $\pi^i: \mathbb{R}^m \to \mathbb{R}$ 是将空间 $\mathbb{R}^m$ 的每个点 $x = (x^1, \cdots, x^m)$ 与它的第 i 个坐标 x^i 相对应的映射, 即

$$\pi^i(x) = x^i.$$

如果 $a = (a^1, \cdots, a^m)$, 那么,

$$\text{当}\, x \to a\, \text{时,}\quad \text{有}\, \pi^i(x) \to a^i.$$

如果 $m > 1$. 当 $x \to \infty$ 时, 函数 $x \mapsto \pi^i(x)$ 既不趋于有限数, 也不趋于无穷远. 同时

$$\text{当}\, x \to \infty\, \text{时,}\quad \text{有}\, f(x) = \sum_{i=1}^{m} (\pi^i(x))^2 \to \infty.$$

不要认为多变量函数的极限可以通过逐次按坐标求极限的方法来得到, 下面的例题可以使人相信这个看法.

例 2 设函数 $f: \mathbb{R}^2 \to \mathbb{R}$ 在点 $(x, y) \in \mathbb{R}^2$ 的定义是

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2}, & x^2 + y^2 \neq 0, \\ 0, & x^2 + y^2 = 0. \end{cases}$$

则 $f(0, y) = f(x, 0) = 0$ 而当 $x \neq 0$, 有 $f(x, x) = \dfrac{1}{2}$.

因此, 这个函数当 $(x,y)\to(0,0)$ 时没有极限.

但是

$$\lim_{y\to0}(\lim_{x\to0}f(x,y))=\lim_{y\to0}(0)=0,$$
$$\lim_{x\to0}(\lim_{y\to0}f(x,y))=\lim_{x\to0}(0)=0.$$

例 3 对于函数

$$f(x,y)=\begin{cases}-\dfrac{x^2-y^2}{x^2+y^2}, & x^2+y^2\neq0,\\ 0, & x^2+y^2=0,\end{cases}$$

有

$$\lim_{x\to0}(\lim_{y\to0}f(x,y))=\lim_{x\to0}\left(-\frac{x^2}{x^2}\right)=-1,$$
$$\lim_{y\to0}(\lim_{x\to0}f(x,y))=\lim_{y\to0}\left(+\frac{y^2}{y^2}\right)=+1.$$

例 4 对于函数

$$f(x,y)=\begin{cases}x+y\sin\dfrac{1}{x}, & x\neq0,\\ 0, & x=0,\end{cases}$$

有

$$\lim_{(x,y)\to(0,0)}f(x,y)=0,$$
$$\lim_{x\to0}(\lim_{y\to0}f(x,y))=0,$$

但是, 累次极限

$$\lim_{y\to0}(\lim_{x\to0}f(x,y))$$

根本就不存在.

例 5 函数

$$f(x,y)=\begin{cases}\dfrac{x^2y}{x^4+y^2}, & x^2+y^2\neq0,\\ 0, & x^2+y^2=0,\end{cases}$$

当点 (x,y) 沿任意射线 $x=\alpha t, y=\beta t$ 趋于坐标原点时以零为极限.

但是, 函数在形如 (a,a^2)(其中 $a\neq0$) 的任意点处的值等于 $\dfrac{1}{2}$, 因此当 $(x,y)\to(0,0)$ 时, 函数没有极限.

2. 多变量连续函数及其性质 设 E 是空间 $\mathbb{R}^m$ 中的集合, $f: E \to \mathbb{R}^n$ 是定义在 E 上在 $\mathbb{R}^n$ 中取值的函数.

定义 6 函数 $f: E \to \mathbb{R}^n$ 称为在点 $a \in E$ 是连续的, 如果对于函数值 $f(a)$ 的任意邻域$V(f(a))$, 存在点 a 在 E 中的邻域$U_E(a)$, 使得 $f(U_E(a)) \subset V(f(a))$.

也就是

$(f: E \to \mathbb{R}^n$ 在 $a \in E$ 连续$):=(\forall V(f(a))\ \exists U_E(a)$

$$(f(U_E(a)) \subset V(f(a)))).$$

我们看出, 定义 6 在形式上与第四章 §1 的定义 1 中我们熟悉的实值函数的连续性是完全一样的. 和那里一样, 我们可以把上述定义改写为下面形式

$$(f: E \to \mathbb{R}^n \text{在} a \in E \text{连续}) := (\forall \varepsilon > 0\ \ \exists \delta > 0\ \forall x \in E\ (d(x,a) < \delta \Rightarrow d(f(x), f(a)) < \varepsilon)).$$

或者, 当点 a 是 E 的极限点时, 也可写成

$$(f: E \to \mathbb{R}^n \text{在} a \in E \text{连续}) := (\lim_{E \ni x \to a} f(x) = f(a)).$$

正如在第四章中已经指出的那样, 连续概念中使我们感兴趣的是点 $a \in E$ 为 E 的极限点且函数在点 a 有定义的情形.

从定义 6 及关系式 (2) 得到, 用关系式

$$(x^1, \cdots, x^m) = x \stackrel{f}{\mapsto} y = (y^1, \cdots, y^n) = (f^1(x^1, \cdots, x^m), \cdots, f^n(x^1, \cdots, x^m))$$

给出的映射 $f: E \to \mathbb{R}^n$ 在某点连续当且仅当每个函数 $y^i = f^i(x^1, \cdots, x^m)$ $(i = 1, \cdots, n)$ 在这点连续.

特别地, 当 $I \subset \mathbb{R}$ 是区间, 函数 $f^1(x), \cdots, f^n(x)$ 在 I 上连续时, 称由

$$x \mapsto y = (y^1, \cdots, y^n) = (f^1(x), \cdots, f^n(x))$$

定义的映射 $f: I \to \mathbb{R}^n$ 为 $\mathbb{R}^n$ 中的道路. 这就是说, $\mathbb{R}^n$ 中的道路是从实轴上的区间 $I \subset \mathbb{R}$ 到 $\mathbb{R}^n$ 中的连续映射.

类似于实值函数在一点的振幅定义, 我们引入函数值在 $\mathbb{R}^n$ 中的函数在一点的振幅概念.

设 $E \subset \mathbb{R}^m, a \in E, B_E(a;r) := E \cap B(a;r)$.

定义 7 量

$$\omega(f;a) := \lim_{r \to +0} \omega(f; B_E(a;r))$$

称为函数 $f: E \to \mathbb{R}^n$ 在点 $a \in E$ 的振幅.

从连续函数的定义 6, 并注意柯西准则及极限性质, 我们可以得到经常用到的连续函数局部性质, 列举如下:

连续函数的局部性质

a) 映射$f: E \to \mathbb{R}^n$ 在点 $a \in E \subset \mathbb{R}^m$ 连续当且仅当 $\omega(f; a) = 0$.

b) 在点 $a \in E$ 连续的映射 $f: E \to \mathbb{R}^n$ 在此点的某个邻域$U_E(a)$ 上有界.

c) 若集合 $Y \subset \mathbb{R}^n$ 上的映射$g: Y \to \mathbb{R}^k$ 在点 $y_0 \in Y$ 连续, 集合 $X \subset \mathbb{R}^m$ 上的映射 $f: X \to Y$ 在点 $x_0 \in X$ 连续, 且 $f(x_0) = y_0$, 则映射 $g \circ f: X \to \mathbb{R}^k$ 存在且在点 $x_0 \in X$ 连续.

对于实值函数, 除上述性质外还有下面的性质

d) 若函数 $f: E \to \mathbb{R}$ 在点 $a \in E$ 连续且 $f(a) > 0 (f(a) < 0)$, 则存在点 a 在 E 中的邻域$U_E(a)$, 使得对一切 $x \in U_E(a)$, 有 $f(x) > 0 (f(x) < 0)$.

e) 若函数 $f: E \to \mathbb{R}$ 和函数 $g: E \to \mathbb{R}$ 在点 $a \in E$ 连续, 则它们的线性组合 $(\alpha f + \beta g): E \to \mathbb{R}$, 其中 $\alpha, \beta \in \mathbb{R}$, 乘积 $(f \cdot g): E \to \mathbb{R}$, 商 $\left(\dfrac{f}{g}\right): E \to \mathbb{R}$, 其中 $g(x) \neq 0, x \in E$ 在点 $a \in E$ 连续.

我们规定, 如果函数 $f: E \to \mathbb{R}^n$ 在 E 的每一点连续, 则称函数 f 在 E 上连续.

集合 E 上一切连续函数 $f: E \to \mathbb{R}^n$ 的集合记作 $C(E; \mathbb{R}^n)$. 如果按上下文意思函数值域很清楚, 则可简记作 $C(E)$. 通常这种省略记号是用在 $\mathbb{R}^n = \mathbb{R}$ 的情形.

例 6 从 $\mathbb{R}^m$ 到 $\mathbb{R}$ 的函数 $(x^1, \cdots, x^m) \xrightarrow{\pi^i} x^i \quad (i = 1, \cdots, m)$ (称为投影) 显然在任一点 $a = (a^1, \cdots, a^m) \in \mathbb{R}^m$ 连续, 这是因为 $\lim\limits_{x \to a} \pi^i(x) = a^i = \pi^i(a)$.

例 7 对于任何定义在 $\mathbb{R}$ 上的函数 $x \mapsto f(x)$, 例如 $x \mapsto \sin x$, 可以考察定义在 $\mathbb{R}^2$ 上的函数 $(x, y) \xmapsto{F} f(x)$. 这时, 如果函数 f 在 $\mathbb{R}$ 上连续, 那么, 新函数 $(x, y) \xmapsto{F} f(x)$ 在 $\mathbb{R}^2$ 上连续, 这可以直接由连续的定义来验证, 也可以由函数 F 是连续函数的复合函数 $(f \circ \pi^1)(x, y)$ 来证实.

特别地, 由此, 并注意 c) ,e), 得到, 例如函数

$$f(x, y) = \sin x + e^{xy}, f(x, y) = \operatorname{arctg}(\ln(|x| + |y| + 1))$$

在 $\mathbb{R}^2$ 上连续.

注意, 上述论证本质上是局部的, 而例子中的函数 f 与 F 分别在整个实轴 $\mathbb{R}$ 或平面 $\mathbb{R}^2$ 上定义, 这种情形是偶然的.

例 8 例 2 中的函数 $f(x, y)$ 在空间 $\mathbb{R}^2$ 中除点 $(0, 0)$ 外在任一点连续. 注意, 虽然函数 $f(x, y)$ 在点 $(0, 0)$ 不连续, 但这个函数当其中一个变量固定时, 对另一个变量而言是连续的.

例 9 若函数 $f: E \to \mathbb{R}^n$ 在集合 E 上连续, 而 $\tilde{E}$ 是 E 的子集, 那么, 函数 f 在这个子集上的限制 $f|_{\tilde{E}}$ 在 $\tilde{E}$ 上连续, 这可以由函数在一点的连续性定义直接得到证实.

现在转入连续函数的全局性质的研究. 首先给出几个定义.

定义 8 从集合 $E \subset \mathbb{R}^m$ 到 $\mathbb{R}^n$ 中的映射 $f: E \to \mathbb{R}^n$ 称为在 E 上的**一致连续映射**, 如果对于任意的 $\varepsilon > 0$, 存在 $\delta > 0$, 对任意的使得 $d(x_1, x_2) < \delta$ 的 x_1、$x_2 \in E$, 有 $d(f(x_1), f(x_2)) < \varepsilon$.

这里, $d(x_1, x_2)$ 与 $d(f(x_1), f(x_2))$ 分别是 $\mathbb{R}^m$ 与 $\mathbb{R}^n$ 中两点间的距离.

当 $m = n = 1$ 时, 上述定义就是我们熟悉的数值函数一致连续性的定义.

定义 9 集合 $E \subset \mathbb{R}^m$ 称为**路连通的**, 如果对于 E 中任意一对点 x_0 与 x_1, 总存在一条以 x_0 与 x_1 为两个端点的道路 $\Gamma: I \to E$, 并且它的承载子在 E 中.

换句话说, 从 E 中任一点 x_0 出发不超出 E 的范围可以达到 E 中另外的任何一个点 x_1.

因为我们暂时不考察除路连通集以外的其他连通集概念, 所以为了简便, 我们暂时称路连通集为**连通集**.

定义 10 称空间 $\mathbb{R}^m$ 中的开连通集为 $\mathbb{R}^m$ 中的**区域**.

例 10 $\mathbb{R}^m$ 中的球 $B(a;r)(r>0)$ 是区域. 我们已经知道 $B(a;r)$ 是 $\mathbb{R}^m$ 中的开集, 现在证明: 球是连通的. 设 $x_0 = (x_0^1, \cdots, x_0^m)$, $x_1 = (x_1^1, \cdots, x_1^m)$ 是球的两个点, 用函数 $x^i(t) = tx_1^i + (1-t)x_0^i \quad (i = 1, \cdots, m)$ 表示定义在闭区间 $0 \leqslant t \leqslant 1$ 上连接 x_0 与 x_1 的道路, 而且它的承载子在球中, 这是因为, 由闵可夫斯基不等式, 对任意的 $t \in [0,1]$, 有

$$
\begin{aligned}
d(x(t), a) = \sqrt{\sum_{i=1}^{m}(x^i(t) - a^i)^2} &= \sqrt{\sum_{i=1}^{m}(t(x_1^i - a^i) + (1-t)(x_0^i - a^i))^2} \\
&\leqslant \sqrt{\sum_{i=1}^{m}(t(x_1^i - a^i))^2} + \sqrt{\sum_{i=1}^{m}((1-t)(x_0^i - a^i))^2} \\
&= t\sqrt{\sum_{i=1}^{m}(x_1^i - a^i)^2} + (1-t)\sqrt{\sum_{i=1}^{m}(x_0^i - a^i)^2} \\
&< tr + (1-t)r = r.
\end{aligned}
$$

例 11 用方程 $(x^1)^2 + (x^2)^2 = r^2 \quad (r > 0)$ 给出的圆周 (一维球面) 是 $\mathbb{R}^2$ 中的子集. 设 $x^1 = r\cos t, x^2 = r\sin t$, 我们看到, 圆周上任一对点都可用沿圆周的道路联结, 也就是说, 圆周是连通集. 但是, 这个集在 $\mathbb{R}^2$ 中不是区域, 因为它在 $\mathbb{R}^2$ 中不是

开集.

现在叙述基本的

连续函数的整体性质

a) 如果映射 $f: K \to \mathbb{R}^n$ 在紧集 $K \subset \mathbb{R}^m$ 上连续, 则它在 K 上一致连续.

b) 如果映射 $f: K \to \mathbb{R}^n$ 在紧集 $K \subset \mathbb{R}^m$ 上连续, 则它在 K 上有界.

c) 如果函数 $f: K \to \mathbb{R}$ 在紧集 $K \subset \mathbb{R}^m$ 上连续, 则它在 K 上具有最大、最小值.

d) 如果函数 $f: E \to \mathbb{R}$ 在连通集 E 上连续, 且 f 在 $a, b \in E$ 具有值 $f(a) = A, f(b) = B$, 则对 A 与 B 之间任意的数 C, 存在点 $c \in E$, 使得 $f(c) = C$.

在第四章 §2 中, 我们已经研究了单变量数值函数的局部与整体性质, 并且给出了它们的证明, 这些证明, 除 d) 以外, 都可转移到上述更一般的情形, 唯一需要改动的是把表达式 $|x_1 - x_2|$ 与 $|f(x_1) - f(x_2)|$ 分别改为 $d(x_1, x_2)$ 与 $d(f(x_1), f(x_2))$, 这里 d 是所考察的点所在空间中的距离. 下面我们给出 d) 的证明.

◀ d) 设 $\Gamma: I \to E$ 是 E 中的道路, $I = [\alpha, \beta] \subset \mathbb{R}, \Gamma(\alpha) = a, \Gamma(\beta) = b, \Gamma$ 是 I 上的连续映射. 由于 E 是连通集, 这样的道路是存在的. 函数 $f \circ \Gamma: I \to \mathbb{R}$ 作为连续函数的复合函数是连续的, 所以, 存在点 $\gamma \in [\alpha, \beta]$, 使得 $f \circ \Gamma(\gamma) = C$. 令 $c = \Gamma(\gamma)$, 则有 $c \in E$ 且 $f(c) = C$. ▶

例 12 在 $\mathbb{R}^m$ 中用方程

$$(x^1)^2 + \cdots + (x^m)^2 = r^2$$

给出的球面 $S(0; r)$ 是紧集.

事实上, 从函数

$$(x^1, \cdots, x^m) \mapsto \sum_{i=1}^{m} (x^i)^2$$

的连续性得知球面是闭集, 而从球面上的点都满足 $|x^i| \leqslant r \quad (i = 1, \cdots, m)$ 得知它是有界的.

函数

$$(x^1, \cdots, x^m) \mapsto (x^1)^2 + \cdots + (x^k)^2 - (x^{k+1})^2 - \cdots - (x^m)^2$$

在整个空间 $\mathbb{R}^m$ 上连续, 因此, 它在球面上的限制连续, 由连续函数的整体性质 c) 知道它在球面上取到最小值与最大值. 在球面上的两点 $(r, 0, \cdots, 0)$ 与 $(0, \cdots, 0, r)$ 处考察这个函数, 它分别取值 r^2 与 $-r^2$. 因为球面是连通集(参看本节末的习题 3), 根据连续函数的整体性质 d) 可以断定, 在球面上存在一点, 使得上述函数在此点的值为零.

例 13 开集 $\mathbb{R}^m \setminus S(0;r)(r>0)$ 不是区域, 因为它不是连通集. 事实上, 如果 $\Gamma: I \to \mathbb{R}^m$ 是道路, 它的一个端点是 $x_0=(0,\cdots,0)$, 另一个端点是 $x_1=(x_1^1,\cdots,x_1^m)$, 且 $(x_1^1)^2+\cdots+(x_1^m)^2>r^2$, 设函数 $f:\mathbb{R}^m \to \mathbb{R}$ 由关系式 $(x^1,\cdots,x^m)\overset{f}{\mapsto}(x^1)^2+\cdots+(x^m)^2$ 确定, 则 Γ 与 f 的复合映射是区间 I 上的连续函数, 而且它在两个端点处的值分别小于与大于 r^2, 因此在闭区间 I 上存在一点 γ, 使得 $(f\circ\Gamma)(\gamma)=r^2$. 这样, 道路 Γ 的承载子上的点 $x_\gamma=\Gamma(\gamma)$ 应处在球面 $S(0;r)$ 上. 我们证明了, 从球 $B(0;r)\subset\mathbb{R}^m$ 中一点出发不穿过它的边界 $S(0;r)$ 是走不出去的.

练 习

1. 设 $f\in C(\mathbb{R}^m;\mathbb{R})$. 证明:

a) 集合 $E_1=\{x\in\mathbb{R}^m|f(x)<c\}$ 是 $\mathbb{R}^m$ 中的开集.

b) 集合 $E_2=\{x\in\mathbb{R}^m|f(x)\leqslant c\}$ 是 $\mathbb{R}^m$ 中的闭集.

c) 集合 $E_3=\{x\in\mathbb{R}^m|f(x)=c\}$ 是 $\mathbb{R}^m$ 中的闭集.

d) 如果 $x\to\infty$ 时, $f(x)\to+\infty$, 则 E_2, E_3 是 $\mathbb{R}^m$ 中的紧集.

e) 对任意的函数 $f:\mathbb{R}^m\to\mathbb{R}$, 集合 $E_4=\{x\in\mathbb{R}^m|\omega(f,x)\geqslant\varepsilon\}$ 是 $\mathbb{R}^m$ 中的闭集.

2. 证明: 映射 $f:\mathbb{R}^m\to\mathbb{R}^n$ 连续当且仅当 $\mathbb{R}^n$ 中任一开集的原像是 $\mathbb{R}^m$ 中的开集.

3. 证明:

a) 连通集 E 在连续映射 $f:E\to\mathbb{R}^n$ 下的像 $f(E)$ 是连通集.

b) 有公共点的连通集的并集是连通集.

c) 半球面 $(x^1)^2+\cdots+(x^m)^2=1, x^m\geqslant 0$, 是连通集.

d) 球面 $(x^1)^2+\cdots+(x^m)^2=1$ 是连通集.

e) 若 $E\subset\mathbb{R}$ 是连通集, 则 E 是 $\mathbb{R}$ 上的区间 (即闭区间, 半开区间, 开区间, 射线或整个数轴).

f) 若 x_0 与 x_1 分别是集合 $M\subset\mathbb{R}^m$ 的内点与外点, 则以 x_0 与 x_1 为端点的任何一条道路的承载子与集合 M 的边界相交.

第八章　多变量函数微分学

§1. $\mathbb{R}^m$ 中的线性结构

1. 作为向量空间的 $\mathbb{R}^m$　从代数教程中已经很熟悉向量空间的概念了.

如果在 $\mathbb{R}^m$ 中对于元素 $x_1=(x_1^1,\cdots,x_1^m)$ 与 $x_2=(x_2^1,\cdots,x_2^m)$, 按照公式

$$x_1+x_2=(x_1^1+x_2^1,\cdots,x_1^m+x_2^m) \tag{1}$$

引进加法, 对于 $\mathbb{R}^m$ 中的元素 $x=(x^1,\cdots,x^m)$ 与 $\mathbb{R}$ 中的数 $\lambda\in\mathbb{R}$, 按照公式

$$\lambda x=(\lambda x^1,\cdots,\lambda x^m) \tag{2}$$

引进数乘法, 那么 $\mathbb{R}^m$ 便成为实数域上的线性空间. 它的点现在可以称为向量. 向量

$$e_i=(0,\cdots,0,1,0,\cdots,0)\quad(i=1,\cdots,m) \tag{3}$$

(这里数 1 仅在第 i 个位置出现) 组成 $\mathbb{R}^m$ 的极大线性无关组, 因此称 $\mathbb{R}^m$ 为m 维向量空间.

$\mathbb{R}^m$ 中任一向量 x 可以按照基底 (3) 展开, 即表成形式

$$x=x^1e_1+\cdots+x^me_m. \tag{4}$$

我们约定, 向量的指标记在向量的下方, 向量的坐标的指标则记在它的上方, 使用它有很多方便的地方. 特别是按照爱因斯坦[①]约定可把 (4) 式简写成如下形式

$$x=x^ie_i, \tag{5}$$

[①]爱因斯坦 (A. Einstein)(1879—1955) 是 20 世纪最伟大的物理学家. 他在量子学方面, 特别是相对论方面的工作给整个现代物理学以革命性的影响.

这里认为同一个指标在上方与下方都出现，就意味着关于这个指标在它的变化范围内求和.

2. 线性映射 $L:\mathbb{R}^m\to\mathbb{R}^n$ 我们知道，从一个向量空间 X 到另一个向量空间 Y 的映射 $L:X\to Y$ 称为线性的，如果对于任意的 $x_1,x_2\in X$ 和 $\lambda_1,\lambda_2\in\mathbb{R}$，有

$$L(\lambda_1x_1+\lambda_2x_2)=\lambda_1L(x_1)+\lambda_2L(x_2).$$

使我们感兴趣的是线性映射$L:\mathbb{R}^m\to\mathbb{R}^n$.

如果 $\{e_1,\cdots,e_m\}$ 与 $\{\tilde{e}_1,\cdots,\tilde{e}_n\}$ 分别是 $\mathbb{R}^m$ 与 $\mathbb{R}^n$ 中给定的基底，那么，在知道了展开式

$$L(e_i)=a_i^1\tilde{e}_1+\cdots+a_i^n\tilde{e}_n=a_i^j\tilde{e}_j\quad(i=1,\cdots,m)\tag{6}$$

即基底向量在线性映射$L:\mathbb{R}^m\to\mathbb{R}^n$ 下的像以后，由映射 L 的线性性质，对于 $\mathbb{R}^m$ 中任一向量

$$h=h^1e_1+h^2e_2+\cdots+h^me_m=h^ie_i,$$

可以求出它在 L 下的像 $L(h)$ 在 $\mathbb{R}^n$ 的基底 $\{\tilde{e}_1,\cdots,\tilde{e}_n\}$ 下的展开式

$$L(h)=L(h^ie_i)=h^iL(e_i)=h^ia_i^j\tilde{e}_j=a_i^jh^i\tilde{e}_j\tag{7}$$

写成坐标形式是

$$L(h)=(a_i^1h^i,\cdots,a_i^nh^i).\tag{8}$$

在给定 $\mathbb{R}^n$ 的基底后，可以把映射 $L:\mathbb{R}^m\to\mathbb{R}^n$ 看作是由 n 个 (坐标) 映射 $L^j:\mathbb{R}^m\to\mathbb{R}$ 构成的序组

$$L=(L^1,\cdots,L^n).\tag{9}$$

考虑到 (8) 式，容易断定映射$L:\mathbb{R}^m\to\mathbb{R}^n$ 是线性的当且仅当 (9) 式中每个映射 L^j 都是线性的.

如果把 (9) 式写成列的形式，那么由 (8) 式得

$$L(h)=\begin{pmatrix}L^1(h)\\ \vdots\\ L^n(h)\end{pmatrix}=\begin{pmatrix}a_1^1&\cdots&a_m^1\\ \vdots&&\vdots\\ a_1^n&\cdots&a_m^n\end{pmatrix}\begin{pmatrix}h^1\\ \vdots\\ h^m\end{pmatrix}.\tag{10}$$

这样，在 $\mathbb{R}^m$ 与 $\mathbb{R}^n$ 中给定基底以后，可以建立线性映射$L:\mathbb{R}^m\to\mathbb{R}^n$ 与 $m\times n$ 矩阵 $(a_i^j)(i=1,\cdots,m;j=1,\cdots,n)$ 之间的 1–1 对应. 线性映射L 的矩阵 (a_i^j) 的第 i 列元素是向量 $e_i\in\{e_1,\cdots,e_m\}$ 在 L 下的像 $L(e_i)$ 的坐标. 任一向量 $h=h^ie_i\in\mathbb{R}^m$ 在 L 下的像 $L(h)$ 的坐标可以用线性映射的矩阵与向量 h 的坐标构成的列向量的乘积 (10) 得到.

如果在 $\mathbb{R}^n$ 中有向量空间结构, 那么可以引入关于映射 $f_1: X \to \mathbb{R}^n$ 与映射 $f_2: X \to \mathbb{R}^n$ 的线性组合 $\lambda_1 f_1 + \lambda_2 f_2$ 的概念, 它用下式定义

$$(\lambda_1 f_1 + \lambda_2 f_2)(x) := \lambda_1 f_1(x) + \lambda_2 f_2(x). \tag{11}$$

特别地, 线性映射$L_1: \mathbb{R}^m \to \mathbb{R}^n$ 与 $L_2: \mathbb{R}^m \to \mathbb{R}^n$ 的线性组合是映射

$$h \mapsto \lambda_1 L_1(h) + \lambda_2 L_2(h) = L(h),$$

显然, 它是线性的, 这个映射的矩阵是对应的映射L_1 与 L_2 的矩阵的线性组合.

线性映射 $A: \mathbb{R}^m \to \mathbb{R}^n$ 与 $B: \mathbb{R}^n \to \mathbb{R}^k$ 的复合 $C = B \circ A$ 显然是线性映射, 它的矩阵由 (10) 式知道, 是用映射B 的矩阵左乘映射A 的矩阵得到的矩阵. 顺便指出, 我们已经知道的矩阵乘法规则, 就是为了适应线性映射的复合映射所对应的矩阵而作的规定.

3. $\mathbb{R}^m$ 中的范数. 数值

$$\|x\| = \sqrt{(x^1)^2 + \cdots + (x^m)^2} \tag{12}$$

称为 $\mathbb{R}^m$ 中向量 $x = (x^1, \cdots, x^m)$ 的范数.

由这个定义及闵可夫斯基不等式得到:

1° $\|x\| \geqslant 0$.

2° $(\|x\| = 0) \Leftrightarrow (x = 0)$.

3° $\|\lambda x\| = |\lambda|\|x\|$, 其中 $\lambda \in \mathbb{R}$.

4° $\|x_1 + x_2\| \leqslant \|x_1\| + \|x_2\|$.

一般地, 向量空间 X 上任何一个满足条件 1°—4° 的函数 $\|\ \ \|: X \to \mathbb{R}$ 称为向量空间 X 的范数. 有时为了明确所讲的向量范数是哪一个空间的范数, 在范数记号右下角标记空间的记号, 例如可以写成 $\|x\|_{\mathbb{R}^m}$ 或 $\|x\|_{\mathbb{R}^n}$. 但是, 如果关于空间及其范数按照上下文的意思很清楚的话, 我们照例是不这样做的.

由 (12) 式知

$$\|x_1 - x_2\| = d(x_1, x_2). \tag{13}$$

其中 $d(x_1, x_2)$ 是度量空间 $\mathbb{R}^m$ 中点 x_1 与 x_2 的距离.

从 (13) 式看出, 当 $x \to x_0$, 条件 $d(x, x_0) \to 0$ 与条件 $\|x - x_0\| \to 0$ 等价.

特别地, 由 (13) 式知

$$\|x\| = d(0, x).$$

范数的性质 4° 称为三角形不等式, 现在已经很清楚这样称呼的理由了.

三角形不等式按照数学归纳法可以推广到任意有限个向量相加的情形, 即有下面的不等式

$$\|x_1 + \cdots + x_k\| \leqslant \|x_1\| + \cdots + \|x_k\|.$$

向量范数的存在, 使得能够对于函数 $f: X \to \mathbb{R}^m$ 与 $g: X \to \mathbb{R}^n$ 在空间 $\mathbb{R}^m$ 与 $\mathbb{R}^n$ 中的值进行比较.

设 $\mathcal{B}$ 是 X 中的基, 如果 $\|f(x)\|_{\mathbb{R}^m} = o(\|g(x)\|_{\mathbb{R}^n})$ 在 $\mathcal{B}$ 下成立, 则记作 $f(x) = o(g(x))$ 或对于基 $\mathcal{B}, f = o(g)$.

如果把映射 $f: X \to \mathbb{R}^m$ 写成坐标形式 $f(x) = (f^1(x), \cdots, f^m(x))$, 那么由不等式

$$|f^i(x)| \leqslant \|f(x)\| \leqslant \sum_{i=1}^{m} |f^i(x)|, \tag{14}$$

可以得到下面有用的命题:

$$(\text{对于基}\mathcal{B}, f = o(g)) \Leftrightarrow (\text{对于基}\mathcal{B}, f^i = o(g), i = 1, \cdots, m). \tag{15}$$

还约定, 如果 $\|f(x)\|_{\mathbb{R}^m} = O(\|g(x)\|_{\mathbb{R}^n})$ 在 $\mathcal{B}$ 下成立, 则记作: 对于基 $\mathcal{B}, f = O(g)$.

这样, 由 (14) 得到

$$(\text{对于基}\mathcal{B}, f = O(g)) \Leftrightarrow (\text{对于基}\mathcal{B}, f^i = O(g), i = 1, \cdots, m). \tag{16}$$

例 设 $L: \mathbb{R}^m \to \mathbb{R}^n$ 是线性映射, $h = x^1 e_1 + \cdots + x^m e_m$ 是 $\mathbb{R}^m$ 中任一向量, 我们来估计 $\|L(h)\|_{\mathbb{R}^n}$.

$$\|L(h)\|_{\mathbb{R}^n} = \left\|\sum_{i=1}^{m} h^i L(e_i)\right\| \leqslant \sum_{i=1}^{m} \|L(e_i)\|\ |h^i| \leqslant \left(\sum_{i=1}^{m} \|L(e_i)\|\right) \|h\|. \tag{17}$$

由此可以断定, 有

$$L(h) = O(h), \text{当} h \to 0 \text{时}, \tag{18}$$

特别地, 从上式可以得到, 当 $x \to x_0$ 时, 有

$$L(x - x_0) = L(x) - L(x_0) \to 0.$$

即线性映射$L: \mathbb{R}^m \to \mathbb{R}^n$ 在任一点 $x_0 \in \mathbb{R}^m$ 都连续. 从 (17) 式还可看出, 线性映射是一致连续的.

4. $\mathbb{R}^m$ 的欧几里得结构 从代数知道实向量空间中数量积的概念, 它是对于空间的向量对定义的一个数值函数 $\langle x, y\rangle$, 并且具有性质:

$$\langle x, x\rangle \geqslant 0.$$

$$(\langle x, x\rangle = 0) \Leftrightarrow (x = 0).$$

$$\langle x_1, x_2\rangle = \langle x_2, x_1\rangle.$$

$$\langle \lambda x_1, x_2\rangle = \lambda\langle x_1, x_2\rangle, \text{其中} \lambda \in \mathbb{R}.$$

$$\langle x_1 + x_2, x_3\rangle = \langle x_1, x_3\rangle + \langle x_2, x_3\rangle.$$

特别地, 从这些性质得到, 如果在空间 $\mathbb{R}^m$ 中确定了基底$\{e_1, \cdots, e_m\}$, 那么用向量 $x = (x^1, \cdots, x^m)$ 与 $y = (y^1, \cdots, y^m)$ 的坐标可以把数量积$\langle x, y\rangle$ 写成双线性形式

$$\langle x, y\rangle = g_{ij}x^i y^j, \tag{19}$$

上式右端的简写式指的是关于指标 i 与 j 的求和式, 其中 $g_{ij} = \langle e_i, e_j\rangle$.

两向量称为*正交的*, 如果它们的数量积为零.

基底 $\{e_1, \cdots, e_m\}$ 称为*正交规范的*, 如果 $g_{ij} = \delta_{ij}$, 其中

$$\delta_{ij} = \begin{cases} 0, i \neq j, \\ 1, i = j. \end{cases}$$

若 $\mathbb{R}^m$ 的基底为正交规范基底, 则向量 x 与 y 的数量积有很简单的形式:

$$\langle x, y\rangle = \delta_{ij}x^i y^j,$$

或

$$\langle x, y\rangle = x^1 y^1 + \cdots + x^m y^m. \tag{20}$$

数量积具有这种形式的坐标称为*笛卡儿坐标*.

我们知道, 规定了数量积的空间 $\mathbb{R}^m$ 称为欧几里得空间.

数量积 (20) 式与向量范数 (12) 式之间有明显的关系:

$$\langle x, x\rangle = \|x\|^2.$$

从代数知道, 有下面的不等式:

$$\langle x, y\rangle^2 \leqslant \langle x, x\rangle\langle y, y\rangle.$$

特别地, 由此可以证明, 对任意一对向量, 存在一个角 $\varphi \in [0, \pi]$, 使得

$$\langle x, y\rangle = \|x\|\ \|y\| \cos\varphi.$$

这个角称为*向量 x 与 y 之间的夹角*. 正是由于这个原因, 当两个向量的数量积为零时, 我们自然地认为它们是正交的.

下面我们叙述在代数中已经知道的下述简单但很有用的重要的事实, 对于我们也是有益的.

在欧几里得空间中, 任何一个线性函数 $L : \mathbb{R}^m \to \mathbb{R}$ 都能表达成

$$L(x) = \langle \xi, x\rangle.$$

的形式, 其中 ξ 是由函数 L 唯一确定的 $\mathbb{R}^m$ 中的向量.

§2. 多变量函数的微分

1. 多变量可微函数和函数在一点的微分

定义 1 定义在集合 $E\subset\mathbb{R}^m$ 上的函数 $f:E\to\mathbb{R}^n$ 称为在极限点 $x\in E$ 是可微的, 如果在 E 上有

$$\boxed{f(x+h)-f(x)=L(x)h+\alpha(x;h),} \tag{1}$$

其中 $L(x):\mathbb{R}^m\to\mathbb{R}^n$ 是关于 h 的线性函数[①], 当 $h\to 0, x+h\in E$ 时, $\alpha(x;h)=o(h)$.

向量

$$\Delta x(h):=(x+h)-x=h,$$
$$\Delta f(x;h):=f(x+h)-f(x)$$

分别称为*自变量增量*与*函数增量*(与这个自变量增量相对应), 按照传统习惯, 分别简记作 Δx 与 $\Delta f(x)$ (它们都是 h 的函数).

关系式 (1) 中的线性函数 $L(x):\mathbb{R}^m\to\mathbb{R}^n$ 称为函数 $f:E\to\mathbb{R}^n$ 在点 $x\in E$ 的*微分, 切映射或导映射*.

函数 $f:E\to\mathbb{R}^n$ 在点 $x\in E$ 的微分记作 $df(x)$、$Df(x)$ 或 $f'(x)$.

根据引进的这些记号, 可将 (1) 式写成

$$f(x+h)-f(x)=f'(x)h+\alpha(x;h),$$

或

$$\Delta f(x;h)=df(x)h+\alpha(x;h).$$

我们注意到微分实际上是对从点 $x\in\mathbb{R}^m$ 出发的位移 h 定义的. 为了强调这点, 我们把一个向量空间 $\mathbb{R}^m$ 与点 $x\in\mathbb{R}^m$ 联系起来, 记作 $T\mathbb{R}^m(x)$ 或 $T\mathbb{R}^m_x$, 向量空间 $T\mathbb{R}^m_x$ 可以解释为紧贴在点 $x\in\mathbb{R}^m$ 上的向量全体所成的集合, 我们称 $T\mathbb{R}^m_x$ 为 $\mathbb{R}^m$ 在点 $x\in\mathbb{R}^m$ 的*切空间*. 这个术语的来源以后再作解释.

微分在向量 $h\in T\mathbb{R}^m_x$ 的值是紧贴在点 $f(x)$ 上的向量 $f'(x)h\in T\mathbb{R}^n_{f(x)}$, 这个向量可以逼近由自变量 x 的增量引起的函数增量 $f(x+h)-f(x)$.

于是, $df(x)$ 或 $f'(x)$ 是线性映射 $f'(x):T\mathbb{R}^m_x\to T\mathbb{R}^n_{f(x)}$.

我们看出, 在一点可微的多变量向量值函数与我们已研究过的单变量可微函数在下述性质上是完全一样的, 即如果它在这点的增量 $\Delta f(x;h)$ 作为自变量增量 h 的函数, 在精确到一个比 $h\to 0$ 高阶的无穷小量 $\alpha(x;h)$ 的条件下, 是线性的.

①同一维情况一样, 我们允许用 $L(x)h$ 代替 $L(x)(h)$, 同时指出, 定义中的 $\mathbb{R}^m$ 与 $\mathbb{R}^n$ 都装备了 §1 中所规定的范数

2. 实值函数的偏导数与微分 如果把 $\mathbb{R}^n$ 中的向量 $f(x+h), f(x), L(x)h, \alpha(x;h)$ 写成坐标形式, 那么等式 (1) 等价于下面 n 个实值函数的等式:

$$f^i(x+h) - f^i(x) = L^i(x)h + \alpha^i(x;h) \quad (i = 1, \cdots, n), \tag{2}$$

由 §1 的关系式 (9) 与 (15) 知, $L^i(x): \mathbb{R}^m \to \mathbb{R}$ 是线性函数, 而当 $h \to 0, x+h \in E$ 时, 有 $\alpha^i(x;h) = o(h), (i = 1, \cdots, n)$.

因此, 下面命题是正确的.

命题 1 *集合 $E \subset \mathbb{R}^m$ 上的映射 $f: E \to \mathbb{R}^n$ 在极限点 $x \in E$ 可微当且仅当映射 f 的每个坐标函数 $f^i: E \to \mathbb{R} (i = 1, \cdots, n)$ 在点 $x \in E$ 可微.*

由于关系式 (1) 与 (2) 是等价的, 因此为了求映射 $f: E \to \mathbb{R}^n$ 的微分 $L(x)$, 只要求出它的坐标函数 $f^i: E \to \mathbb{R}$ 的微分 $L^i(x)$ 就可以了.

现在我们考虑定义在集合 $E \subset \mathbb{R}^m$ 上的实值函数 $f: E \to \mathbb{R}$, 且设它在 E 的内点 $x \in E$ 可微. 下面将主要研究 E 是 $\mathbb{R}^m$ 中的区域的情形. 如果 x 是 E 的内点, 那么对于点 x 的任意足够小的位移 h, 点 $x+h$ 同样属于 E, 也就是说, 也在 $f: E \to \mathbb{R}$ 的定义域 E 中.

如果用坐标形式表示点 $x = (x^1, \cdots, x^m)$, 向量 $h = (h^1, \cdots, h^m)$ 和线性函数 $L(x)h = a_1(x)h^1 + \cdots + a_m(x)h^m$, 那么条件

$$f(x+h) - f(x) = L(x)h + o(h), \text{当} h \to 0 \text{时}, \tag{3}$$

可改写为

$$\begin{aligned} &f(x^1 + h^1, \cdots, x^m + h^m) - f(x^1, \cdots, x^m) \\ &= a_1(x)h^1 + \cdots + a_m(x)h^m + o(h), \text{当} h \to 0 \text{时}, \end{aligned} \tag{4}$$

其中 $a_1(x), \cdots, a_m(x)$ 是依赖于点 x 的实数.

我们现在来求这些实数. 为此用特殊的位移

$$h_i = h^i e_i = 0 \cdot e_1 + \cdots + 0 \cdot e_{i-1} + h^i e_i + 0 \cdot e_{i+1} + \cdots + 0 \cdot e_m$$

代替任意的位移 h, 显然, h_i 与 $\mathbb{R}^m$ 的基底 $\{e_1, \cdots, e_m\}$ 中的向量 e_i 共线.

当 $h = h_i$ 时, 显然有 $\|h\| = |h^i|$, 因此, 从 (4) 式得到

$$\begin{aligned} &f(x^1, \cdots, x^{i-1}, x^i + h^i, x^{i+1}, \cdots, x^m) - f(x^1, \cdots, x^i, \cdots, x^m) \\ &= a_i(x)h^i + o(h^i), \quad \text{当} h^i \to 0 \text{时}. \end{aligned} \tag{5}$$

由此看出, 如果函数 $f(x^1, \cdots, x^m)$ 除第 i 个自变量外, 其余自变量皆固定不变, 那么, 这时得出第 i 个变量的函数在点 x^i 是可微的.

这样一来, 从等式 (5) 得到

$$a_i(x) = \lim_{h^i \to 0} \frac{f(x^1, \cdots, x^{i-1}, x^i + h^i, x^{i+1}, \cdots, x^m) - f(x^1, \cdots, x^i, \cdots, x^m)}{h^i}. \tag{6}$$

定义 2 极限 (6) 称为函数 $f(x)$ 在点 $x = (x^1, \cdots, x^m)$ 关于变量 x^i 的偏导数. 下面每个符号都可用来表示它:

$$\frac{\partial f}{\partial x^i}(x), \partial_i f(x), D_i f(x), f'_{x^i}(x).$$

例 1 设 $f(u,v) = u^3 + v^2 \sin u$, 则

$$\partial_1 f(u,v) = \frac{\partial f}{\partial u}(u,v) = 3u^2 + v^2 \cos u,$$

$$\partial_2 f(u,v) = \frac{\partial f}{\partial v}(u,v) = 2v \sin u.$$

例 2 设 $f(x,y,z) = \operatorname{arctg}(xy^2) + e^z$, 则

$$\partial_1 f(x,y,z) = \frac{\partial f}{\partial x}(x,y,z) = \frac{y^2}{1 + x^2y^4},$$

$$\partial_2 f(x,y,z) = \frac{\partial f}{\partial y}(x,y,z) = \frac{2xy}{1 + x^2y^4},$$

$$\partial_3 f(x,y,z) = \frac{\partial f}{\partial z}(x,y,z) = e^z.$$

这样, 我们证明了

命题 2 *如果定义在集合 $E \subset \mathbb{R}^m$ 上的函数 $f: E \to \mathbb{R}$ 在 E 的内点 x 可微, 那么函数 f 在点 x 关于每个变量有偏导数, 并且它的微分由这些偏导数唯一确定:*

$$df(x)h = \frac{\partial f}{\partial x^1}(x)h^1 + \cdots + \frac{\partial f}{\partial x^m}(x)h^m, \tag{7}$$

公式 (7) 按照爱因斯坦约定可以简写为

$$df(x)h = \partial_i f(x)h^i. \tag{8}$$

例 3 假如我们知道 (很快就会知道的) 例 2 中的函数 $f(x,y,z)$ 在点 $(0,1,0)$ 可微, 那么立刻可以写出

$$df(0,1,0)h = 1 \cdot h^1 + 0 \cdot h^2 + 1 \cdot h^3 = h^1 + h^3.$$

由此有

$$f(h^1, 1+h^2, h^3) - f(0,1,0) = df(0,1,0)h + o(h) \quad (h \to 0)$$

或

$$\operatorname{arctg}(h^1(1+h^2)^2) + e^{h^3} = 1 + h^1 + h^3 + o(h) \quad (h \to 0).$$

例 4 函数 $x=(x^1,\cdots,x^m)\xrightarrow{\pi^i}x^i$ 把点 $x\in\mathbb{R}^m$ 对应于它的第 i 个坐标. 对于这个函数有

$$\Delta\pi^i(x,h)=(x^i+h^i)-x^i=h^i,$$

也就是说, 这个函数的增量本身是 h 的线性函数 $h\xrightarrow{\pi^i}h^i$, 因此 $\Delta\pi^i(x,h)=d\pi^i(x)h$. 由于对任意的点 $x\in\mathbb{R}^m$ 有 $d\pi^i(x)h=h^i$, 在这个意义上映射 $d\pi^i(x)=d\pi^i$ 实际上是不依赖于点 x 的. 如果把 $\pi^i(x)$ 记作 $x^i(x)$, 则有 $d\pi^i(x)h=dx^ih=h^i$.

注意到这种状况以及公式 (8), 现在我们可以把任何函数的微分表示成它的自变量 $x\in\mathbb{R}^m$ 的坐标的微分的线性组合, 亦即有

$$df(x)=\partial_i f(x)dx^i=\frac{\partial f}{\partial x^1}(x)dx^1+\cdots+\frac{\partial f}{\partial x^m}(x)dx^m, \tag{9}$$

这是因为, 对于任意向量 $h\in T\mathbb{R}^m_x$ 有

$$df(x)h=\partial_i f(x)h^i=\partial_i f(x)dx^ih.$$

3. 映射的微分的坐标表示. 雅可比矩阵 我们已经得到实值函数 $f:E\to\mathbb{R}$ 的微分公式 (7). 而由此, 根据关系式 (1) 与 (2) 的等价性, 对于任意的定义在集合 $E\subset\mathbb{R}^m$ 上的映射 $f:E\to\mathbb{R}^n$, 如果它在 E 的内点 $x\in E$ 可微, 那么就可以把它在点 x 的微分 $df(x)$ 表成坐标形式:

$$\begin{aligned}df(x)h&=\begin{pmatrix}df^1(x)h\\ \vdots\\ df^n(x)h\end{pmatrix}=\begin{pmatrix}\partial_i f^1(x)h^i\\ \vdots\\ \partial_i f^n(x)h^i\end{pmatrix}\\ &=\begin{pmatrix}\frac{\partial f^1}{\partial x^1}(x)&\cdots&\frac{\partial f^1}{\partial x^m}(x)\\ \vdots&&\vdots\\ \frac{\partial f^n}{\partial x^1}(x)&\cdots&\frac{\partial f^n}{\partial x^m}(x)\end{pmatrix}\begin{pmatrix}h^1\\ \vdots\\ h^m\end{pmatrix}.\end{aligned}\tag{10}$$

定义 3 设映射 $f:E\to\mathbb{R}^n$, 在点 $x\in E\subset\mathbb{R}^m$ 可微, f 的坐标函数在点 x 的偏导数组成的矩阵 $(\partial_i f^j(x))(i=1,\cdots,m;j=1,\cdots,n)$ 称为映射 f 在点 x 的雅可比矩阵[①]或 Jacobian[②].

当 $n=1$ 时, 公式 (10) 归结为公式 (7), 当 $n=1,m=1$ 时, 公式 (10) 归结为一个实变量的实值函数的微分.

由关系式 (1) 与 (2) 的等价性及实值函数微分的唯一性得到

---

①雅可比 (J. Jacobi)(1804—1851) —— 著名的德国数学家.

②又常称这个矩阵的行列式 (当它是方阵时) 为 Jacobian.

命题 3 如果定义在集合 $E \subset \mathbb{R}^m$ 上的映射 $f: E \to \mathbb{R}^n$ 在 E 的内点 $x \in E$ 可微, 则它在点 x 的微分是唯一的, 并且映射 $df(x): T\mathbb{R}^m_x \to T\mathbb{R}^n_{f(x)}$ 的坐标表示用 (10) 式给出.

4. 函数在一点的连续性、偏导数和可微性 本段将完成函数在一点可微性的讨论, 指出函数在一点的连续性、偏导数的存在性与可微性之间的相互关系.

在 §1 中 ((17) 式与 (18) 式) 我们得知, 如果 $L: \mathbb{R}^m \to \mathbb{R}^n$ 是线性映射, 那么当 $h \to 0$ 时, $Lh \to 0$, 这样从 (1) 式可以推知, 若函数在一点可微, 则它在此点连续, 这是因为

当 $h \to 0, x+h \in E$ 时, $f(x+h) - f(x) = L(x)h + o(h)$.

当然上述命题的逆命题是不正确的, 这是因为, 我们知道, 在单变量情形这已经就是错误的了.

这样, 函数在一点的可微性与连续性的关系在高维情形和一维情形是一样的.

但是微分与偏导数之间的关系完全是另一种样子. 在一维情形, 即一个实变量的实值函数情形, 函数在一点存在微分与存在导数是等价的. 对于多变量函数, 我们已经证明 (命题 2) 函数在定义域内点的可微性保证了它在此点关于每个变量的偏导数的存在性, 但是相反的命题是不对的.

例 5 函数

$$f(x^1, x^2) = \begin{cases} 0, x^1x^2 = 0, \\ 1, x^1x^2 \neq 0. \end{cases}$$

在坐标轴上等于零, 因此在点 (0,0) 处, 它的两个偏导数是

$$\partial_1 f(0,0) = \lim_{h^1 \to 0} \frac{f(h^1,0) - f(0,0)}{h^1} = \lim_{h^1 \to 0} \frac{0-0}{h^1} = 0,$$
$$\partial_2 f(0,0) = \lim_{h^2 \to 0} \frac{f(0,h^2) - f(0,0)}{h^2} = \lim_{h^2 \to 0} \frac{0-0}{h^2} = 0.$$

但是这个函数在点 (0,0) 不可微, 因为显然它在此点间断.

例 5 引进的这个函数, 除了原点 (0,0) 外, 在数轴上任意点处, 它的任何一个偏导数都不存在. 但是函数

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2}, & x^2+y^2 \neq 0, \\ 0, & x^2+y^2 = 0. \end{cases}$$

(即第七章 §2 例 2 中的函数) 在平面上所有点 (x,y) 处存在偏导数, 但是它在坐标原点间断, 因而它在此点不可微.

这样, 虽然可以正确地写出 (7),(8) 式的右边部分, 但仍不能保证它是我们的函数的微分, 这是因为函数可能原本就不是可微的.

要不是发现了 (这将在以后证明) 函数偏导数的连续性足以保证函数的可微性, 这种状况将会成为多变量函数微分计算的严重障碍.

§3. 微分法的基本定律

1. 微分法运算的线性性质

定理 1 如果定义在集合 $E \subset \mathbb{R}^m$ 上的映射 $f_1 : E \to \mathbb{R}^n, f_2 : E \to \mathbb{R}^n$ 在点 $x \in E$ 可微, 则它们的线性组合 $(\lambda_1 f_1 + \lambda_2 f_2) : E \to \mathbb{R}^n$ 在点 x 可微, 且有

$$(\lambda_1 f_1 + \lambda_2 f_2)'(x) = (\lambda_1 f_1' + \lambda_2 f_2')(x). \tag{1}$$

等式 (1) 指出了微分运算, 即在一点一个映射对应于它在该点的微分的运算, 是一切在此点可微的映射 $f : E \to \mathbb{R}^n$ 组成的向量空间上的线性运算. 等式 (1) 的左边 $(\lambda_1 f_1 + \lambda_2 f_2)'(x)$ 按定义是线性映射, 而右边 $(\lambda_1 f_1' + \lambda_2 f_2')(x)$ 是线性映射$f_1'(x) : \mathbb{R}^m \to \mathbb{R}^n$ 与 $f_2'(x) : \mathbb{R}^m \to \mathbb{R}^n$ 的线性组合, 由 §1 知它也是线性映射. 定理 1 断言这两个映射完全一样.

$$\begin{aligned}
◀ \ & (\lambda_1 f_1 + \lambda_2 f_2)(x+h) - (\lambda_1 f_1 + \lambda_2 f_2)(x) \\
&= (\lambda_1 f_1(x+h) + \lambda_2 f_2(x+h)) - (\lambda_1 f_1(x) + \lambda_2 f_2(x)) \\
&= \lambda_1 (f_1(x+h) - f_1(x)) + \lambda_2 (f_2(x+h) - f_2(x)) \\
&= \lambda_1 (f_1'(x)h + o(h)) + \lambda_2 (f_2'(x)h + o(h)) \\
&= (\lambda_1 f_1'(x) + \lambda_2 f_2'(x))h + o(h). \quad ▶
\end{aligned}$$

如果考察的是实值函数, 那么关于乘法运算与除法运算 (当分母不为零) 有下面的

定理 2 如果定义在集合 $E \subset \mathbb{R}^m$ 上的函数 $f : E \to \mathbb{R}, g : E \to \mathbb{R}$ 在点 $x \in E$ 可微, 则

a) 它们的乘积在点 x 可微, 且

$$(f \cdot g)'(x) = g(x)f'(x) + f(x)g'(x). \tag{2}$$

b) 当 $g(x) \neq 0$ 时, 它们的商在点 x 可微, 且

$$\left(\frac{f}{g}\right)'(x) = \frac{1}{g^2(x)}(g(x)f'(x) - f(x)g'(x)). \tag{3}$$

这个定理的证明同第五章,§2, 定理 1 的证明完全一样, 我们不去重复它了.

关系式 (1)、(2)、(3) 可以用另外的微分记号改写成下面的形式：

$$d(\lambda_1 f_1 + \lambda_2 f_2)(x) = (\lambda_1 df_1 + \lambda_2 df_2)(x),$$

$$d(f \cdot g)(x) = g(x)df(x) + f(x)dg(x),$$

$$d\left(\frac{f}{g}\right)(x) = \frac{1}{g^2(x)}(g(x)df(x) - f(x)dg(x)).$$

我们来考察如何用映射的坐标形式表示这些等式. 我们知道, 如果映射$\varphi : E \to \mathbb{R}^n$ 在 $E \subset \mathbb{R}^m$ 的内点 x 可微, 把它写成坐标形式

$$\varphi(x) = \begin{pmatrix} \varphi^1(x^1, \cdots, x^m) \\ \vdots \\ \varphi^n(x^1, \cdots, x^m) \end{pmatrix},$$

那么与它在点 x 的微分 $d\varphi(x) : \mathbb{R}^m \to \mathbb{R}^n$ 相对应的是雅可比矩阵

$$\varphi'(x) = \begin{pmatrix} \partial_1\varphi^1 & \cdots & \partial_m\varphi^1 \\ \vdots & & \vdots \\ \partial_1\varphi^n & \cdots & \partial_m\varphi^n \end{pmatrix}(x) = (\partial_i\varphi^j)(x).$$

在 $\mathbb{R}^m$ 与 $\mathbb{R}^n$ 中给定基底的情况下, 线性映射$L : \mathbb{R}^m \to \mathbb{R}^n$ 与它的 $m \times n$ 矩阵之间存在 1—1 对应, 因此可以把线性映射L 与它的矩阵看成一回事.

对于雅可比矩阵的记号, 我们通常使用的是 $f'(x)$ 而不是 $df(x)$, 这是为了更符合一维情形下导数与微分概念的传统区别.

这样, 根据在集合 E 的内点 x 处的微分的唯一性, 我们得到用雅可比矩阵表示的 (1),(2),(3) 式的如下坐标形式:

$$(\partial_i(\lambda_1 f_1^j + \lambda_2 f_2^j))(x) = (\lambda_1\partial_i f_1^j + \lambda_2\partial_i f_2^j)(x)$$
$$(i = 1, \cdots, m; j = 1, \cdots, n), \tag{1'}$$

$$(\partial_i(f \cdot g))(x) = g(x)\partial_i f(x) + f(x)\partial_i g(x)$$
$$(i = 1, \cdots, m), \tag{2'}$$

$$\left(\partial_i\left(\frac{f}{g}\right)\right)(x) = \frac{1}{g^2(x)}(g(x)\partial_i f(x) - f(x)\partial_i g(x))$$
$$(i = 1, \cdots, m). \tag{3'}$$

例如, 从上述矩阵的分量等式可以推出, 实值函数 $f(x^1, \cdots, x^m)$ 与 $g(x^1, \cdots, x^m)$ 的乘积关于变量 x^i 的偏导数应是:

$$\frac{\partial(f \cdot g)}{\partial x^i}(x^1, \cdots, x^m) = g(x^1, \cdots, x^m)\frac{\partial f}{\partial x^i}(x^1, \cdots, x^m)$$
$$+ f(x^1, \cdots, x^m)\frac{\partial g}{\partial x^i}(x^1, \cdots, x^m).$$

我们注意到, 无论是这个等式还是矩阵等式 (1′), (2′), (3′) 都是偏导数定义以及一个实变量的实值函数的微分法则的显然的推论. 但是, 我们知道偏导数的存在性还不能断定多变量函数的可微性. 因此, 除了重要的和完全显然的等式 (1′), (2′), (3′) 之外, 在定理 1 与定理 2 中, 关于映射的微分的存在性假定也是一个特殊的角色.

最后我们指出, 利用归纳法可将等式 (2) 推广为有限个可微实值函数乘积 $(f_1 \cdot \cdots \cdot f_k)$ 的微分式:

$$d(f_1 \cdot \cdots \cdot f_k)(x) = (f_2 \cdot \cdots \cdot f_k)(x)df_1(x) + \cdots + (f_1 \cdot \cdots \cdot f_{k-1})(x)df_k(x).$$

2. 复合映射的微分法

a. 基本定理

定理 3 如果集合 $X \subset \mathbb{R}^m$ 到集合 $Y \subset \mathbb{R}^n$ 的映射 $f: X \to Y$ 在点 $x \in X$ 可微, 而映射 $g: Y \to \mathbb{R}^k$ 在 $y = f(x) \in Y$ 可微, 则复合映射 $g \circ f: X \to \mathbb{R}^k$ 在点 x 可微, 且复合映射的微分 $d(g \circ f)(x): T\mathbb{R}^m_x \to T\mathbb{R}^k_{g(f(x))}$ 等于微分 $df(x): T\mathbb{R}^m_x \to T\mathbb{R}^n_y$ 与微分 $dg(y): T\mathbb{R}^n_y \to T\mathbb{R}^k_{g(y)}$ 的复合 $dg(y) \circ df(x)$.

这个定理的证明, 几乎完全重复第五章 §2 的定理 2 的证明. 为了把注意力集中于现在出现的新情况, 我们再次把证明的全过程写出来, 但是不准备详细讨论我们已经熟悉的技术性问题.

◀ 利用映射 f 与 g 在点 x 与 $y = f(x)$ 的可微性和微分 $g'(x)$ 的线性, 可得

$$\begin{aligned}(g \circ f)(x+h) - (g \circ f)(x) &= g(f(x+h)) - g(f(x)) \\ &= g'(f(x))(f(x+h) - f(x)) + o(f(x+h) - f(x)) \\ &= g'(y)(f'(x)h + o(h)) + o(f(x+h) - f(x)) \\ &= g'(y)(f'(x)h) + g'(y)(o(h)) + o(f(x+h) - f(x)) \\ &= (g'(y) \circ f'(x))h + \alpha(x;h).\end{aligned}$$

其中 $g'(y) \circ f'(x)$ 是线性映射(因为它是线性映射的复合), 而

$$\alpha(x;h) = g'(y)(o(h)) + o(f(x+h) - f(x)).$$

但是像证明 §1 关系式 (17),(18) 那样, 得

$$g'(y)(o(h)) = o(h) \quad (h \to 0),$$
$$f(x+h) - f(x) = f'(x)h + o(h) = O(h) + o(h) = O(h) \quad (h \to 0)$$

和

$$o(f(x+h) - f(x)) = o(O(h)) = o(h) \quad (h \to 0)$$

因此

$$\alpha(x;h) = o(h) + o(h) = o(h) \quad (h \to 0),$$

从而, 定理得证. ►

现在我们把定理 3 写成坐标形式, 如果 x 是集合 X 的内点, 且

$$f'(x) = \begin{pmatrix} \partial_1 f^1(x) & \cdots & \partial_m f^1(x) \\ \vdots & & \vdots \\ \partial_1 f^n(x) & \cdots & \partial_m f^n(x) \end{pmatrix} = (\partial_i f^j)(x),$$

而 $y = f(x)$ 是集合 Y 的内点, 且

$$g'(y) = \begin{pmatrix} \partial_1 g^1(y) & \cdots & \partial_n g^1(y) \\ \vdots & & \vdots \\ \partial_1 g^k(y) & \cdots & \partial_n g^k(y) \end{pmatrix} = (\partial_j g^l)(y),$$

那么

$$\begin{aligned} (g \circ f)'(x) &= \begin{pmatrix} \partial_1(g^1 \circ f)(x) & \cdots & \partial_n(g^1 \circ f)(x) \\ \vdots & & \vdots \\ \partial_1(g^k \circ f)(x) & \cdots & \partial_n(g^k \circ f)(x) \end{pmatrix} = (\partial_i(g^l \circ f))(x) \\ &= \begin{pmatrix} \partial_1 g^1(y) & \cdots & \partial_n g^1(y) \\ \vdots & & \vdots \\ \partial_1 g^k(y) & \cdots & \partial_n g^k(y) \end{pmatrix} \begin{pmatrix} \partial_1 f^1(x) & \cdots & \partial_m f^1(x) \\ \vdots & & \vdots \\ \partial_1 f^n(x) & \cdots & \partial_m f^n(x) \end{pmatrix} \\ &= (\partial_j g^l(y) \cdot \partial_i f^j(x)). \end{aligned}$$

在等式

$$(\partial_i(g^l \circ f))(x) = (\partial_j g^l(f(x)) \cdot \partial_i f^j(x)) \tag{4}$$

右边矩阵中的元素 $\partial_j g^l(f(x)) \cdot \partial_i f^j(x)$ 是对指标 j 从 1 到 n 求和的简写记号.

与等式 $(1')$, $(2')$, $(3')$ 不同, 甚至从有关矩阵的分量等式方面看, 关系式 (4) 也不是无需证明的.

我们考察定理的某些重要的特殊情形.

b. 复合实值函数的偏导数与微分　设 $z = g(y^1, \cdots, y^n)$ 是实变量 $y^1, \cdots, y^n$ 的实值函数, 而 $y^1, \cdots, y^n$ 中的每一个又是变量 $x^1, \cdots, x^m$ 的函数 $y^j = f^j(x^1, \cdots, x^m)$ $(j = 1, \cdots, n)$. 设函数 g 与 $f^j (j = 1, \cdots, n)$ 可微, 我们来求 $f: X \to Y$ 与 $g: Y \to \mathbb{R}$ 的复合映射的偏导数 $\dfrac{\partial(g \circ f)}{\partial x^i}(x)$.

根据公式 (4), 按我们的条件, 其中的 $l=1$, 得

$$\partial_i(g\circ f)(x)=\partial_j g(f(x))\cdot\partial_i f^j(x) \tag{5}$$

或者更详细地写作

$$\begin{aligned}\frac{\partial z}{\partial x^i}(x)&=\frac{\partial(g\circ f)}{\partial x^i}(x^1,\cdots,x^m)=\frac{\partial g}{\partial y^1}\frac{\partial y^1}{\partial x^i}+\cdots+\frac{\partial g}{\partial y^n}\frac{\partial y^n}{\partial x^i}\\&=\partial_1 g(f(x))\cdot\partial_i f^1(x)+\cdots+\partial_n g(f(x))\cdot\partial_i f^n(x),(i=1,\cdots,m).\end{aligned}$$

c. 函数在一点沿向量的导数和梯度 我们来考察在空间 $\mathbb{R}^3$ 的某个区域 G 中液体或气体的定常流, 术语《定常》指的是流体在区域 G 中每一点的速度都不随时间而改变, 当然在区域 G 中不同的点可以具有不同的速度. 譬如, $f(x)=f(x^1,x^2,x^3)$ 是流体在点 $x=(x^1,x^2,x^3)\in G$ 的压力. 如果我们转入按规律 $x=x(t)$(t 是时间) 的流动, 那么在时刻 t 的压力值是 $f\circ x(t)=f(x(t))$, 而压力沿流动轨道关于时间的变化速度显然是导数 $\frac{d(f\circ x)}{dt}(t)$. 在 $f(x^1,x^2,x^3)$ 是 G 中的可微函数的假定下, 我们将能算出它. 按照复合函数的微分法则, 有

$$\frac{d(f\circ x)}{dt}(t)=\frac{\partial f}{\partial x^1}(x(t))\dot{x}^1(t)+\frac{\partial f}{\partial x^2}(x(t))\dot{x}^2(t)+\frac{\partial f}{\partial \dot{x}^3}(x(t))\dot{x}^3(t), \tag{6}$$

其中 $\dot{x}^i(t)=\frac{dx^i}{dt}(t)(i=1,2,3)$.

因为 $(\dot{x}^1,\dot{x}^2,\dot{x}^3)(t)=v(t)$ 是流体位移在时刻 t 的速度向量, 而 $(\partial_1 f,\partial_2 f,\partial_3 f)(x)$ 是函数 f 在点 x 的微分 $df(x)$ 的坐标形式, 因此等式 (6) 也可改写为

$$\frac{d(f\circ x)}{dt}(t)=df(x(t))v(t), \tag{7}$$

就是说, 所求的量是函数 $f(x)$ 在点 $x=x(t)$ 的微分 $df(x(t))$ 在我们的运动的速度向量 $v(t)$ 上的值.

特别地, 如果当 $t=0$ 时, 有 $x_0=x(0)$, 则

$$\frac{d(f\circ x)}{dt}(0)=df(x_0)v, \tag{8}$$

其中 $v=v(0)$ 是速度向量在 $t=0$ 的值.

关系式 (8) 的右方仅仅依赖于点 $x_0\in G$ 和在 x_0 的速度向量 v, 而与轨道 $x=x(t)$ 的具体形式无关, 只要满足条件 $\dot{x}(0)=v$ 就行了. 这就是说, 对任意的形如

$$x(t)=x_0+v(t)+\alpha(t) \tag{9}$$

(其中 $\alpha(t)=o(t)(t\to 0)$) 的轨迹上, 等式 (8) 的左端将有同一数值, 这是因为它完全由给定点 x_0 和此点处的向量 $v\in T\mathbb{R}^3_{x_0}$ 所确定, 特别地, 假如我们想直接计算出等式 (8) 左端的数值 (也就算出了右端的数值), 那么可以取下面的函数作为运动规律

$$x(t)=x_0+vt \tag{10}$$

这是一个速度为 v 的匀速运动, 且在时刻 $t=0$ 有 $x(0)=x_0$.

现在给出下面的

定义 1 如果函数 f 定义在点 $x_0\in\mathbb{R}^m$ 的邻域中, 而 $v\in T\mathbb{R}^m_{x_0}$ 是在点 x_0 的向量, 则称量

$$\boxed{D_vf(x_0):=\lim_{t\to0}\frac{f(x_0+vt)-f(x_0)}{t}} \tag{11}$$

为函数 f 在点 x_0 沿向量 v 的导数(如果上述极限存在的话).

由上面的引言看出, 如果函数 f 在点 x_0 可微, 那么对任意形如 (9), 特别地, 形如 (10) 的函数 $x(t)$, 有

$$D_vf(x_0)=\frac{d(f\circ x)}{dt}(0)=df(x_0)v, \tag{12}$$

其坐标形式为

$$D_vf(x_0)=\frac{\partial f}{\partial x^1}(x_0)v^1+\cdots+\frac{\partial f}{\partial x^m}(x_0)v^m. \tag{13}$$

特别地, 对于基底向量 $e_1=(1,0,\cdots,0),\cdots,e_m=(0,\cdots,0,1)$, 从这个公式得

$$D_{e_i}f(x_0)=\frac{\partial f}{\partial x^i}(x_0),(i=1,\cdots,m).$$

由微分 $df(x_0)$ 的线性, 根据等式 (12) 可以断定, 如果函数 f 在点 x_0 可微, 那么对任意的向量 $v_1,v_2\in T\mathbb{R}^m_{x_0}$ 和任意的实数 $\lambda_1,\lambda_2\in\mathbb{R}$, 函数 f 在点 x_0 沿方向 $(\lambda_1v_1+\lambda_2v_2)\in T\mathbb{R}^m_{x_0}$ 有导数, 并且有

$$D_{\lambda_1v_1+\lambda_2v_2}f(x_0)=\lambda_1D_{v_1}f(x_0)+\lambda_2D_{v_2}f(x_0). \tag{14}$$

如果把空间 $\mathbb{R}^m$ 看成是欧几里得空间, 也就是规定了数量积的向量空间 (参看 §1), 那么对任意的线性函数 $L(v)$ 可以写成由 L 确定的向量 $\xi=\xi(L)$ 与变向量 v 的数量积 $\langle\xi,v\rangle$.

特别地, 存在向量 ξ, 使得

$$df(x_0)v=\langle\xi,v\rangle. \tag{15}$$

定义 2 设函数 f 在点 $x_0\in\mathbb{R}^m$ 可微, 如果向量 $\xi\in T\mathbb{R}^m_{x_0}$ 适合于函数 f 在 x_0 的微分 $df(x_0)$ 所满足的等式 (15), 则称 ξ 为函数 f 在点 x_0 的梯度, 且记作 $\mathrm{grad}f(x_0)$.

于是, 由定义有

$$\boxed{df(x_0)v=\langle\mathrm{grad}\,f(x_0),v\rangle.} \tag{16}$$

如果在 $\mathbb{R}^m$ 中选取笛卡儿坐标系, 那么比较 (12),(13) 与 (16) 式, 可以断定, 在这个坐标系下, 梯度有下面的坐标表示式:

$$\mathrm{grad}\,f(x_0)=\left(\frac{\partial f}{\partial x^1},\cdots,\frac{\partial f}{\partial x^m}\right)(x_0) \tag{17}$$

现在我们阐明向量 $\mathrm{grad}f(x_0)$ 的几何意义.

设 $e\in T\mathbb{R}^m_{x_0}$ 是单位向量. 由 (16) 式知

$$D_e f(x_0)=\|\mathrm{grad}\,f(x_0)\|\cos\varphi, \tag{18}$$

其中 φ 是向量 e 与 $\mathrm{grad}f(x_0)$ 的夹角.

因此, 如果 $\mathrm{grad}f(x_0)\neq 0, e=\|\mathrm{grad}f(x_0)\|^{-1}\,\mathrm{grad}\,f(x_0)$, 那么导数 $D_e f(x_0)$ 具有最大值. 当动点从定点 x_0 沿着方向向量 $\mathrm{grad}f(x_0)$ 移动时, 函数 f 的值的增长速度 (f 的值的单位以 $\mathbb{R}^m$ 的单位长为量度单位) 最快且等于 $\|\mathrm{grad}f(x_0)\|$. 当动点从定点 x_0 沿着上述方向的相反方向移动时, 函数 f 的值最急剧地减少. 当动点从定点 x_0 沿着与方向 $\mathrm{grad}f(x_0)$ 垂直的方向移动时, 函数 f 的值的变化速度等于零.

沿给定方向的单位向量的导数, 通常称为*沿给定方向的导数*.

因为欧几里得空间中的单位向量可用方向余弦表示:

$$e=(\cos\alpha_1,\cdots,\cos\alpha_m),$$

其中 α_i 是向量 e 与笛卡儿坐标系的基底向量 e_i 之间的夹角, 所以

$$D_e f(x_0)=\langle\mathrm{grad}\,f(x_0),e\rangle=\frac{\partial f}{\partial x^1}(x_0)\cos\alpha_1+\cdots+\frac{\partial f}{\partial x^m}(x_0)\cos\alpha_m.$$

我们今后将常常遇到向量 $\mathrm{grad}f(x_0)$, 并且它有很多应用, 例如, 根据以上叙述的梯度的几何性质建立的*梯度方法*, 它是 (利用电子计算机) 寻求多变量函数极值的数值搜索方法 (参看本节末的习题 2).

很多重要的向量场, 例如, 牛顿引力场或库仑静电场都是某个数量函数 (称为这个场的*势*或*位*) 的梯度(参看本节末习题 3).

很多物理定律在其陈述中都利用了向量 $\mathrm{grad}f(x)$, 例如, 连续介质力学中研究的无外力作用的理想液体或气体所满足的关系式

$$\rho a=-\,\mathrm{grad}\,p$$

其中 $a=a(x,t)$ 是流体在点 x 和时刻 t 的加速度, 而 $\rho=\rho(x,t)$ 是介质密度, $p=p(x,t)$ 是压力. 这个关系式与质点动力学中的牛顿基本定律 $ma=F$ 相当 (参看本节末习题 4).

关于向量 $\mathrm{grad}f$ 我们以后在向量分析与初等场论中还要研究.

3. 逆映射的微分法

定理 4 *设 $f:U(x)\to V(y)$ 是把点 x 的邻域 $U(x)\subset\mathbb{R}^m$ 映到点 $y=f(x)$ 的邻域 $V(y)\subset\mathbb{R}^m$ 上的映射, 它在点 x 连续且存在逆映射 $f^{-1}:V(y)\to U(x)$, 而且逆映射在点 y 也连续.*

如果映射 f 在点 x 可微且它在点 x 的切映射 $f'(x):T\mathbb{R}^m_x\to T\mathbb{R}^m_y$ 有逆映射 $[f'(x)]^{-1}:T\mathbb{R}^m_y\to T\mathbb{R}^m_x$, 则映射 $f^{-1}:V(y)\to U(x)$ 在点 $y=f(x)$ 可微, 且

$$(f^{-1})'(y)=[f'(x)]^{-1}.$$

因此, 互逆的可微映射在对应的点处有互逆的切映射.

◀ 设

$$f(x) = y, f(x+h) = y+t, t = f(x+h) - f(x);$$

则

$$f^{-1}(y) = x, f^{-1}(y+t) = x+h, h = f^{-1}(y+t) - f^{-1}(y).$$

假定 h 足够小, 使得 $x+h \in U(x)$, 从而, $y+t \in V(y)$. 由 f 在点 x 和 f^{-1} 在点 y 的连续性得到

$$\text{当} h \to 0 \text{时}, t = f(x+h) - f(x) \to 0, \tag{1}$$

$$\text{当} t \to 0 \text{时}, h = f^{-1}(y+t) - f^{-1}(y) \to 0. \tag{2}$$

由 f 在点 x 的可微性得到

$$t = f'(x)h + o(h)(h \to 0). \tag{3}$$

由此可以断定 $t = O(h)(h \to 0)$.(参看 §1 的 (17)、(18) 式.)

我们来证明, 如果 $f'(x)$ 有可逆的线性映射, 则 $h = O(t)(t \to 0)$.

事实上, 由 (3) 式得到

$$[f'(x)]^{-1}t = h + [f'(x)]^{-1}[o(h)](h \to 0), \tag{4}$$

$$[f'(x)]^{-1}t = h + o(h)(h \to 0),$$

$$\|[f'(x)]^{-1}t\| \geqslant \|h\| - \|o(h)\|(h \to 0),$$

$$\|[f'(x)]^{-1}t\| \geqslant \frac{1}{2}\|h\|, \text{当} \|h\| < \delta \text{时},$$

其中数 $\delta > 0$ 选得, 使当 $\|h\| < \delta$ 时, 有 $\|o(h)\| < \frac{1}{2}\|h\|$ 成立. 考虑到 (2) 式, 有

$$\|h\| \leqslant 2\|[f'(x)]^{-1}t\| = O(\|t\|)(t \to 0)$$

它等价于关系式

$$h = O(t) \quad (t \to 0),$$

特别地, 由此推出

$$o(h) = o(t) \quad (t \to 0).$$

这样, 从 (2) 与 (4) 式就得到

$$h = [f'(x)]^{-1}t + o(t) \quad (t \to 0)$$

或

$$f^{-1}(y+t) - f^{-1}(y) = [f'(x)]^{-1}t + o(t) \quad (t \to 0). \quad ▶$$

由代数知道, 如果线性映射$L:\mathbb{R}^m\to\mathbb{R}^m$ 对应的矩阵是 A, 那么 L 的逆线性映射 $L^{-1}:\mathbb{R}^m\to\mathbb{R}^m$ 对应的矩阵是 A 的逆阵 A^{-1}. 由代数我们还知道如何计算逆矩阵 A^{-1} 的元素, 这就给出了一种构造映射 $(f^{-1})'(y)$ 的直接方法.

注意, 当 $m=1$, 即 $\mathbb{R}^m=\mathbb{R}$ 时, 映射 $f:U(x)\to V(y)$ 在点 x 的雅可比行列式是一个数 $f'(x)$ —— 函数 f 在点 x 的导数, 而线性映射$f'(x):T\mathbb{R}_x\to T\mathbb{R}_y$ 是乘这个数的映射 : $h\mapsto f'(x)h$. 这个线性映射有逆映射当且仅当 $f'(x)\neq 0$, 并且逆映射$[f'(x)]^{-1}:T\mathbb{R}_y\to T\mathbb{R}_x$ 的矩阵同样由一个数 $[f'(x)]^{-1}$ 组成, 即数 $f'(x)$ 的倒数. 这就是说, 定理 4 包括了以前已经证明的反函数求导法则.

练　习

1. a) $\mathbb{R}^m$ 中两条道路$t\mapsto x_1(t)$ 与 $t\mapsto x_2(t)$ 在点 $x_0\in\mathbb{R}^m$ 称为等价的, 如果它们在 $t=0$ 可微, $x_1(0)=x_2(0)=x_0$, 且 $d(x_1(t),x_2(t))=o(t)(t\to 0)$.
   证明: 上面指出的关系实际上是等价关系, 即它们有反身性、对称性和传递性.
   b) 证明: 向量 $v\in T\mathbb{R}^m_{x_0}$ 与在点 x_0 的等价的光滑道路类之间有相互单值的对应.
   c) 把切空间 $T\mathbb{R}^m_{x_0}$ 等同于在点 x_0 的等价的光滑道路类; 引进道路类的加法和数乘法.
   d) 在等价道路类中引进的运算依赖于 $\mathbb{R}^m$ 中的坐标系吗?

2. a) 画出函数 $z=x^2+4y^2$ 的图像, 其中 x,y,z 是 $\mathbb{R}^3$ 中的笛卡儿坐标.
   b) 设 $f:G\to\mathbb{R}$ 是定义在区域 $G\subset\mathbb{R}^m$ 上的数值函数, 如果函数 f 在集合 $E\subset G$ 上仅有一个值 ($f(E)=c$), 精确些说, $E=f^{-1}(c)$, 则称集合 E 是函数 f 的等高集(c-等高面). 在 $\mathbb{R}^2$ 中画出函数 $f(x,y)=x^2+4y^2$ 的等高集的图像.
   c) 求函数 $f(x,y)=x^2+4y^2$ 的梯度, 并且证明: 在任一点 (x,y) 处, 向量 $\mathrm{grad}f$ 与函数 f 过此点的等高线垂直.
   d) 利用上述 a),b),c) 的结果, 在曲面 $z=x^2+4y^2$ 上求出点 (2,1,8) 到它的最低点 (0,0,0) 的最短路径.
   e) 为了寻找函数 $f(x,y)=x^2+4y^2$ 的最小值, 你能提出什么适用于电子计算机的算法?

3. 我们说在空间 $\mathbb{R}^m$ 的区域 G 中给定了一个向量场, 如果每一点 $x\in G$ 都对应于某个向量 $v(x)\in T\mathbb{R}^m_x$. 区域 G 上的向量场 $v(x)$ 称为势场(位场), 如果存在一个数值函数 $U:G\to\mathbb{R}$, 使得 $v(x)=\mathrm{grad}U(x)$. 函数 $U(x)$ 称为场 $v(x)$ 的势(位)(在物理中, 通常称 $-U(x)$ 为势函数, 如果向量场是力场, 则称 $U(x)$ 为力函数).
   a) 设 $f_1(x,y)=x^2+y^2$; $f_2(x,y)=-(x^2+y^2)$; $f_3(x,y)=\mathrm{arc\ tg}\dfrac{x}{y}$ $(y>0)$; $f_4(x,y)=xy$. 在笛卡儿坐标平面上画出上面各个函数的梯度场.
   b) 根据牛顿定律, 位于点 $O\in\mathbb{R}^3$ 的质量为 m 的质点对位于点 $x\in\mathbb{R}^3(x\neq 0)$ 的质量为 1 的质点的引力是 $\boldsymbol{F}=-\dfrac{m}{|\boldsymbol{r}|^3}\boldsymbol{r}$, 其中 $\boldsymbol{r}$ 是向量 $\vec{Ox}$(我们省去了引力常数 G). 证明: 区域 $\mathbb{R}^3\backslash O$ 上的向量场 $\boldsymbol{F}(x)$ 是势场.

c) 证明: 放在点 $(\xi_i,\eta_i,\zeta_i)\quad(i=1,\cdots,n)$ 处质量为 $m_i(i=1,\cdots,n)$ 的质点在这些点的外面所建立的牛顿力场的势函数是

$$U(x,y,z)=\sum_{i=1}^{n}\frac{m_i}{\sqrt{(x-\xi_i)^2+(y-\eta_i)^2+(z-\zeta_i)^2}}.$$

d) 求放在点 $(\xi_i,\eta_i,\zeta_i)(i=1,\cdots,n)$ 处电量为 $q_i(i=1,\cdots,n)$ 的点电荷所形成的库仑静电场场强的势函数.

4. 考察在无外力 (包括重力) 作用时, 空间中的理想不可压缩流体的运动.

设

$$\boldsymbol{v}=\boldsymbol{v}(x,y,z,t)$$
$$\boldsymbol{a}=\boldsymbol{a}(x,y,z,t)$$
$$\rho=\rho(x,y,z,t)$$
$$p=p(x,y,z,t)$$

分别是介质在点 (x,y,z) 与时刻 t 的速度、加速度、密度和压力.

理想流体指的是, 在它的任一点压力不依赖于方向.

a) 从流体中分出一个不大的平行六面体微团, 使得其中一个边平行于向量 $\operatorname{grad} p(x,y,z,t)$ (其中 $\operatorname{grad} p$ 按照空间坐标系选取). 借助于压力落差, 估计作用在这个微团上的力, 并给出这个微团 (认为它是不可压缩的) 加速度的近似公式.

b) 试检验, 你在 a) 中得到的结果与欧拉方程

$$\rho\boldsymbol{a}=-\operatorname{grad} p$$

是否一致.

c) 如果曲线上任一点处的切线方向与这点的速度向量的方向相同, 则称此曲线为*流线*. 运动称为*定常的*, 如果函数 $\boldsymbol{v},\boldsymbol{a},\rho,p$ 不依赖于时间 t. 利用 b) 证明: 在定常的不可压缩的流体中沿着流线运动时, 量 $\frac{1}{2}\|\boldsymbol{v}\|^2+\frac{p}{\rho}$ 是常数, 这个事实称为伯努利定律[①].

d) 如果运动发生在靠近地球表面的重力场, 如何修改 a) 与 b) 中的公式? 证明: 在这种情况有 $\rho\boldsymbol{a}=-\operatorname{grad}(gz+p)$, 且因此在定常的不可压缩流体中, 沿着每一条流线, 量

$$\frac{1}{2}\|\boldsymbol{v}\|^2+gz+\frac{p}{\rho}$$

是常数, 其中 g 是重力加速度, z 是流线上点的高 (从某个零水平面算起).

e) 根据以上结果说明, 为什么承载机翼有特殊的向上凸的断面.

f) 在底半径为 R 的圆柱形杯子中装入密度为 ρ, 高度为 h 的不可压缩的理想流体, 然后把这个杯子绕它的轴以角速度 ω 旋转. 利用流体的不可压缩性质, 求流体在定常状态时, 它的曲面方程 $z=f(x,y)$(参看第 5 章 §1 习题 3).

[①] D 伯努利 (D Bernoulli)(1700—1782) —— 瑞士的学者, 是当时最卓越的数学家与物理学家之一.

g) 根据已求出的曲面方程 $z = f(x, y)$, 写出旋转流体微团中任一点 (x, y, z) 处压力 $p = p(x, y, z)$ 的公式. 试检验, 你求出的公式是否满足 d) 中的方程 $\rho a = -\operatorname{grad}(gz + p)$.

h) 现在你能不能解释, 为什么泡湿的茶叶下沉 (虽不太快!), 而当旋转时, 茶叶并不聚集在杯壁附近, 而是在杯底中心.

5. 函数值计算的误差估计

a) 利用可微函数定义与近似等式 $\Delta f(x; h) \approx df(x)h$, 证明: 设 $f(x) = x^1 \cdot \cdots \cdot x^m$ 是 m 个不为零的因子的乘积, 若 δ_i 是第 i 个因子的相对误差, 则它们乘积的相对误差为 $\delta = \delta(f(x); h) \approx \sum\limits_{i=1}^{m} \delta_i$.

b) 利用等式 $d \ln f(x) = \dfrac{1}{f(x)} df(x)$, 再次推导上题结果, 并证明: 一般的分式

$$\frac{f_1 \cdot \cdots \cdot f_n}{g_1 \cdot \cdots \cdot g_k}(x^1, \cdots, x^m)$$

的相对误差是函数 $f_1, \cdots, f_n, g_1, \cdots, g_k$ 的值的误差的和.

6. 齐次函数和欧拉恒等式 定义在区域 $G \subset \mathbb{R}^m$ 上的函数 $f : G \to \mathbb{R}$ 称为*n 次齐次函数(正齐次)*, 如果对于任意的 $x \in \mathbb{R}^m$ 与 $\lambda \in \mathbb{R}$, 且 $x \in G, \lambda x \in G$, 有 $f(\lambda x) = \lambda^n f(x)(f(\lambda x) = |\lambda|^n f(x))$.

在区域 G 中函数称为*局部 n 次齐次的*, 如果它在区域 G 的每一点的某个邻域中是 n 次齐次的.

a) 证明: 在凸域中, 每个局部齐次函数是齐次函数.

b) 设区域 G 是平面 $\mathbb{R}^2$ 去掉射线

$$L = \{(x, y) \in \mathbb{R}^2 | x = 2 \wedge y \geqslant 0\}$$

的部分, 证明: 函数

$$f(x, y) = \begin{cases} \dfrac{y^4}{x}, & x > 2 \wedge y > 0, \\ y^3, & \text{在区域中其余的点}. \end{cases}$$

在 G 中是局部齐次的, 但在 G 中不是齐次的.

c) 指出下面函数在它们的自然定义域中齐次或正齐次的次数:

$$f_1(x^1, \cdots, x^m) = x^1x^2 + x^2x^3 + \cdots + x^{m-1}x^m,$$

$$f_2(x^1, x^2, x^3, x^4) = \frac{x^1x^2 + x^3x^4}{x^1x^2x^3 + x^2x^3x^4},$$

$$f_3(x^1, \cdots, x^m) = |x^1 \cdot \cdots \cdot x^m|^l.$$

d) 对等式 $f(tx) = t^n f(x)$ 按 t 求微分, 证明: 如果可微函数 $f : G \to \mathbb{R}$ 在区域 G 中是局部 n 次齐次的, 则它在 G 中满足下面的*齐次函数的欧拉恒等式*.

$$x^1 \frac{\partial f}{\partial x^1}(x^1, \cdots, x^m) + \cdots + x^m \frac{\partial f}{\partial x^m}(x^1, \cdots, x^m) \equiv nf(x^1, \cdots, x^m).$$

e) 证明: 如果对于在区域中可微的函数 $f : G \to \mathbb{R}$ 欧拉恒等式成立, 则这个函数在 G 中是局部 n 次齐次的. (提示: 证明函数 $\varphi(t) = t^{-n} f(tx)$ 对任意的 $x \in G$ 有定义并且在 $t = 1$ 的某个邻域中是常数.)

7. 齐次函数和量纲方法

1° *物理量的量纲以及物理量相依关系的特性* 物理定律建立了物理量之间的相互联系, 因此如果对于这些量中间的某些量采取某种度量单位, 那么其他量的度量单位也将表示为这些量的度量单位的确定形式, 这样就产生了这样或那样的度量单位制的基本单位和导出单位.

在国际单位制即 SI 制 (System internatinal) 中, 作为基本的力学度量单位, 长度取米 (m), 质量取千克 (kg), 时间取秒 (s).

用基本度量单位表示的导出度量单位的式子称为它的*量纲*, 这个定义下面将会准确界定.

任何一个力学量的量纲可以用上面指出的基本单位的量纲符号 L, M, T 写成公式的形式, 例如速度, 加速度和力的量纲分别有 $[v] = LT^{-1}$, $[a] = LT^{-2}$, $[F] = MLT^{-2}$. 量纲符号 L, M, T 是麦克斯韦①建议的.

如果物理定律不依赖于选取的度量单位, 那么, 作为这个不变性的反映, 物理量数字表征之间的函数关系:

$$x_0 = f(x_1, \cdots, x_k, x_{k+1}, \cdots, x_n) \tag{*}$$

必具有一定的特点.

例如考察直角三角形斜边与直角边长度之间的关系 $c = f(a, b) = \sqrt{a^2 + b^2}$, 所有边长数值应当是按同样比例变化的, 因此对于任意允许数值 a 与 b, 应当成立关系式 $f(\alpha a, \alpha b) = \varphi(\alpha) f(a, b)$, 并且在此例中 $\varphi(\alpha) = \alpha$.

量纲理论的基本前提 (乍看起来是显然的) 是:*具有物理意义的关系式 (*) 应当是这样的, 当基本单位的尺度改变时, 公式中所有同名量的数值改变同样的倍数.*

特别地, 如果 x_1, x_2, x_3 是基本的独立的物理量, 且函数 $(x_1, x_2, x_3) \mapsto f(x_1, x_2, x_3)$ 值是依赖于它们的第四个物理量, 那么由上述原则, 对任意允许的数值 x_1, x_2, x_3 应当成立等式

$$f(\alpha_1 x_1, \alpha_2 x_2, \alpha_3 x_3) = \varphi(\alpha_1, \alpha_2, \alpha_3) f(x_1, x_2, x_3) \tag{**}$$

其中 φ 是某个具体的函数.

等式 (**) 中的函数 φ 完全描述了所考察的物理量的数值与基本物理量的度量单位尺度变化之间的相依关系. 于是, 这个函数关系应该认为是给定物理量关于确定的基本度量单位的*量纲*

现在给出量纲函数的具体形式.

a) 设 $x \mapsto f(x)$ 是满足条件 $f(\alpha x) = \varphi(\alpha) f(x)$ 的单变量的函数, 其中 f 与 φ 是可微函数. 证明: $\varphi(\alpha) = \alpha^d$.

b) 证明: 等式 (**) 中的量纲函数恒有 $\varphi(\alpha_1, \alpha_2, \alpha_3) = \alpha_1^{d_1} \cdot \alpha_2^{d_2} \cdot \alpha_3^{d_3}$ 的形式, 其中幂的指数 d_1, d_2, d_3 是某些实数. 这样, 譬如确定的基本单位为 L, M, T, 那么幂表达式 $L^{d_1} M^{d_2} T^{d_3}$ 中的指数组 (d_1, d_2, d_3) 同样可以认为是给定物理量的量纲.

c) 在 b) 中已经得到量纲函数有幂函数的形式, 即由每个基本度量单位的幂构成的齐次函数. 如果某个物理量的量纲函数关于某个基本度量单位的齐性次数等于零, 这意味着什

---

①麦克斯韦 (James Clerk Maxwell)(1831—1879) —— 杰出的英国物理学家, 他建立了电磁场的数学理论, 还以气体运动理论、光学理论和力学理论研究而闻名.

么?

2° Π–定理与量纲方法　设 $[x_i] = X_i(i = 0, 1, \cdots, n)$ 是定律 $(*)$ 中物理量的量纲.

假设量 $x_0, x_{k+1}, \cdots, x_n$ 的量纲能够用量 $x_1, \cdots, x_k$ 的量纲表示, 即

$$[x_0] = X_0 = X_1^{p_0^1} \cdot \cdots \cdot X_k^{p_0^k},$$
$$[x_{k+i}] = X_{k+i} = X_1^{p_i^1} \cdot \cdots \cdot X_k^{p_i^k} \quad (i = 1, \cdots, n-k).$$

d) 证明: 同 $(*)$ 式一起有下式成立

$$\alpha_1^{p_0^1} \cdot \cdots \cdot \alpha_k^{p_0^k} x_0 = f(\alpha_1 x_1, \cdots, \alpha_k x_k, \alpha_1^{p_1^1} \cdot \cdots \cdot \alpha_k^{p_1^k} x_{k+1}, \cdots,$$
$$\alpha_1^{p_{n-k}^1} \cdot \cdots \cdot \alpha_k^{p_{n-k}^k} x_{n-k}). \tag{$***$}$$

e) 如果 $x_1, \cdots, x_k$ 相互独立, 那么在 $(***)$ 中令 $\alpha_1 = x_1^{-1}, \cdots, \alpha_k = x_k^{-1}$. 试验证, 这时从 $(***)$ 中得到等式

$$\frac{x_0}{x_1^{p_0^1} \cdot \cdots \cdot x_k^{p_0^k}} = f\left(1, \cdots, 1, \frac{x_{k+1}}{x_1^{p_1^1} \cdot \cdots \cdot x_k^{p_2^k}}, \cdots, \frac{x_n}{x_1^{p_{n-k}^1} \cdot \cdots \cdot x_k^{p_{n-k}^k}}\right),$$

它是无量纲的量 $\Pi, \Pi_1, \cdots, \Pi_{n-k}$ 之间的关系式

$$\Pi = f(1, \cdots, 1, \Pi_1, \cdots, \Pi_{n-k}). \tag{$****$}$$

这样一来, 得到下面的

量纲理论的 Π–定理　*如果在关系式 $(*)$ 中的量 $x_1, \cdots, x_k$ 是相互独立的, 则这个关系式可归结为 $n-k$ 个无量纲的参数的函数 $(****)$.*

f) 证明: 如果 $k = n$, 则根据 Π– 定理, 关系式 $(*)$ 中的函数 f 能够确定到只差一个数因子. 试用上述方法求摆的振动周期表达式 $c(\varphi_0)\sqrt{\dfrac{l}{g}}$(所谓摆, 也就是用长为 l 的线悬挂的一质量为 m 的小球, 它在地球表面振动. φ_0 是初始偏转角).

g) 求出在圆周轨道 (半径为 r) 上旋转, 中心力为 F, 质量为 m 的物体的旋转周期 $P = c\left(\dfrac{mr}{F}\right)^{1/2}$.

h) 对作圆周运动的行星 (卫星), 它们的周期与轨道半径满足开普勒定律$\left(\dfrac{P_1}{P_2}\right)^2 = \left(\dfrac{r_1}{r_2}\right)^3$, 请仿照牛顿, 由它求出万有引力定律$F = G\dfrac{m_1 m_2}{r^\alpha}$ 中幂的指数 α.

§4. 多变量实值函数微分学的基本事实

1. 中值定理

定理 1　*设 $f: G \to \mathbb{R}$ 是定义在区域 $G \subset \mathbb{R}^m$ 上的实值函数, 闭区间 $[x, x+h] \subset G$. 如果函数 f 在闭区间 $[x, x+h]$ 上连续, 在开区间 $]x, x+h[$ 上可微, 则存在点 $\xi \in]x, x+h[$, 使得下式成立*

$$\boxed{f(x+h) - f(x) = f'(\xi)h.} \tag{1}$$

◀ 考察定义在闭区间 [0,1] 上的辅助函数

$$F(t) = f(x+th), \quad 0 \leqslant t \leqslant 1.$$

函数 F 满足拉格朗日定理所有的条件: 它是连续映射的复合, 因而在 [0,1] 上连续, 而作为可微映射的复合, 在]0,1[上可微, 因此存在点 $\theta \in]0,1[$, 使得

$$F(1) - F(0) = F'(\theta) \cdot 1.$$

但是, $F(1) = f(x+h), F(0) = f(x), F'(\theta) = f'(x+\theta h)h$, 由此, 得证定理 1. ▶

现在把关系式 (1) 写成坐标形式.

如果 $x = (x^1, \cdots, x^m), h = (h^1, \cdots, h^m), \xi = (x^1 + \theta h^1, \cdots, x^m + \theta h^m)$, 那么等式 (1) 可写为

$$\begin{aligned} f(x+h) - f(x) &= f(x^1+h^1, \cdots, x^m+h^m) - f(x^1, \cdots, x^m) \\ &= f'(\xi)h = \left(\frac{\partial f}{\partial x^1}(\xi), \cdots, \frac{\partial f}{\partial x^m}(\xi)\right) \begin{pmatrix} h^1 \\ \vdots \\ h^m \end{pmatrix} \\ &= \partial_1 f(\xi) h^1 + \cdots + \partial_m f(\xi) h^m \\ &= \sum_{i=1}^m \partial_i f(x^1 + \theta h^1, \cdots, x^m + \theta h^m) h^i. \end{aligned}$$

利用以前关于重复上、下指标求和的约定, 最后可将上式写成

$$\begin{aligned} &f(x^1+h^1, \cdots, x^m+h^m) - f(x^1, \cdots, x^m) \\ &= \partial_i f(x^1 + \theta h^1, \cdots, x^m + \theta h^m) h^i, \end{aligned} \tag{1'}$$

其中 $0 < \theta < 1$, 且 θ 与 x, h 有关.

注 由于能够找到满足等式 (1) 的某个 "中间的" 点 $\xi \in]x, x+h[$, 因此称定理 1 为中值定理. 在拉格朗日定理 (参看第五章, §3, 第 1 段) 的讨论中我们已经发现, 中值定理是实值函数所特有的. 关于映射的一般的有限增量定理 将在第十章 (第 II 卷) 中给出.

由定理 1 得到有用的

推论 *如果函数 $f: G \to \mathbb{R}$ 在区域 $G \subset \mathbb{R}^m$ 上可微, 并且在 G 的任一点 x 它的微分等于零, 则 f 在 G 上是常数.*

◀ 线性映射等于零等价于它的矩阵的所有元素等于零. 因

$$df(x)h = (\partial_1 f, \cdots, \partial_m f)(x)h,$$

所以对任意的 $x \in G$, 有 $\partial_1 f(x) = \cdots = \partial_m f(x) = 0$.

按定义, 区域是连通开集, 我们将利用这一点.

首先我们证明, 若 $x \in G$, 则存在球 $B(x;r) \subset G$, 使得函数 f 在 $B(x;r)$ 上为常数.

事实上, 若 $x+h \in B(x;r)$, 则 $[x, x+h] \subset B(x;r) \subset G$. 应用关系式 (1) 或 (1′), 得到

$$f(x+h) - f(x) = f'(\xi)h = 0 \cdot h = 0.$$

即 $f(x+h) = f(x)$, 也就是说, f 在球 $B(x;r)$ 上的值皆等于 f 在球心 x 处的值.

现在设 $x_0, x_1 \in G$ 是区域 G 中任意两点, 因 G 是连通的, 所以存在一条联结 $x(0) = x_0$ 与 $x(1) = x_1$ 的道路 $t \mapsto x(t) \in G$. 我们假定连续映射 $t \mapsto x(t)$ 定义在闭区间 [0,1] 上, 并设 $B(x_0;r) \subset G$, 因为 $x(0) = x_0$, 且映射 $t \mapsto x(t)$ 连续, 所以存在数 $0 < \delta \leqslant 1$, 当 $0 \leqslant t \leqslant \delta$ 时, 有 $x(t) \in B(x_0;r) \subset G$. 由上所证, 当 $t \in [0,\delta]$ 时, 有 $f \circ x(t) \equiv f(x_0)$. 设

$$l = \sup\{\delta_1 | f \circ x(t) = f(x_0), 0 \leqslant t \leqslant \delta_1 \leqslant 1\},$$

显然有 $0 < l \leqslant 1$, 由函数 $f(x(t))$ 在 [0,1] 上的连续性有 $f(x(l)) = f(x_0)$. 但这时有 $l = 1$. 事实上, 在相反的情况, 即假设 $l < 1$, 则存在球 $B(x(l);r) \subset G$, 使得 f 在此球上有 $f(x) = f(x(l)) = f(x_0)$, 因函数 $t \mapsto x(t)$ 连续, 则存在 $\Delta > 0$, 对一切满足 $l \leqslant t \leqslant l + \Delta$ 的 t, 有 $x(t) \in B(x(l);r)$, 且

$$f \circ x(t) = f(x(l)) = f(x_0), l \leqslant t \leqslant l + \Delta,$$

这与 l 的定义矛盾. 因此证明了, 对一切 $t \in [0,1]$, 有 $f \circ x(t) = f(x_0)$. 特别地, $f \circ x(1) = f(x_1) = f(x_0)$, 从而验明, 对任意两点 $x_0, x_1 \in G$, 函数 $f: G \to \mathbb{R}$ 的值是相同的. ►

2. 多变量函数可微性的充分条件

定理 2 *设 $f: U(x) \to \mathbb{R}$ 是定义在点 $x = (x^1, \cdots, x^m)$ 的邻域 $U(x) \subset \mathbb{R}^m$ 上的函数.*

如果函数 f 在邻域 $U(x)$ 上每一点的所有偏导数 $\dfrac{\partial f}{\partial x^1}, \cdots, \dfrac{\partial f}{\partial x^m}$ 都存在, 且它们在点 x 连续, 则函数 f 在此点可微.

◄ 不失一般性, 可以认为 $U(x)$ 是球 $B(x;r)$. 易知, 若点 $x = (x^1, \cdots, x^m)$ 与点 $x + h = (x^1 + h^1, \cdots, x^m + h^m)$ 属于球 $B(x;r)$, 则点 $(x^1, x^2 + h^2, \cdots, x^m + h^m), \cdots, (x^1, x^2 \cdots, x^{m-1}, x^m + h^m)$ 以及联结它们的线段皆属于 $B(x;r)$. 利用一个

变量的函数的拉格朗日定理有

$$\begin{aligned} f(x+h)-f(x) &= f(x^1+h^1,\cdots,x^m+h^m)-f(x^1,\cdots,x^m) \\ &= f(x^1+h^1,\cdots,x^m+h^m)-f(x^1,x^2+h^2,\cdots,x^m+h^m) \\ &\quad +f(x^1,x^2+h^2,\cdots,x^m+h^m)-f(x^1,x^2,x^3+h^3,\cdots,x^m+h^m) \\ &\quad +f(x^1,x^2,x^3+h^3,\cdots,x^m+h^m)-\cdots+f(x^1,x^2\cdots,x^{m-1}, \\ &\quad x^m+h^m)-f(x^1,x^2,\cdots,x^m) \\ &= \partial_1 f(x^1+\theta^1h^1,x^2+h^2,\cdots,x^m+h^m)h^1+\partial_2 f(x^1,x^2+\theta^2h^2,x^3 \\ &\quad +h^3,\cdots,x^m+h^m)h^2+\cdots+\partial_m f(x^1,\cdots,x^{m-1},x^m+\theta^mh^m)h^m. \end{aligned}$$

到此, 我们仅仅用到了函数在邻域 $U(x)$ 上偏导数的存在性. 下面要利用偏导数在点 x 的连续性. 继续上面的计算, 得到

$$\begin{aligned} f(x+h)-f(x) = \partial_1 f(x^1,\cdots,x^m)h^1+\alpha_1h^1+\partial_2 f(x^1,\cdots,x^m)h^2 \\ +\alpha_2h^2+\cdots+\partial_m f(x^1,\cdots,x^m)h^m+\alpha_mh^m. \end{aligned}$$

其中 $\alpha_1,\cdots,\alpha_m$ 由于偏导数在点 x 的连续性, 它们当 $h\to 0$ 时, 都趋于零.

但这意味着

$$f(x+h)-f(x)=L(x)h+o(h)\quad (h\to 0),$$

其中 $L(x)h=\partial_1 f(x^1,\cdots,x^m)h^1+\cdots+\partial_m f(x^1,\cdots,x^m)h^m$. ►

从定理 2 推知, 如果函数 $f:G\to\mathbb{R}$ 的偏导数在区域 $G\subset\mathbb{R}^m$ 上连续, 那么函数在区域 G 的任一点都可微.

今后我们约定用记号 $C^{(1)}(G;\mathbb{R})$ 或更简单地用 $C^{(1)}(G)$ 表示在区域 G 上具有连续偏导数的函数的集合.

3. 高阶偏导数 如果函数 $f:G\to\mathbb{R}$ 定义在区域 $G\subset\mathbb{R}^m$ 上, 且对每个变量 $x^i(i=1,\cdots,m)$ 有偏导数 $\dfrac{\partial f}{\partial x^i}(x)$, 那么这个偏导数又是一个新函数 $\partial_i f:G\to\mathbb{R}$, 它同样关于某个 x^j 可以有偏导数 $\partial_j(\partial_i f)(x)$.

函数 $\partial_j(\partial_i f):G\to\mathbb{R}$ 称为函数 f 关于变量 x^i,x^j 的二阶导数,并记作

$$\partial_{ji}f(x)\quad 或 \frac{\partial^2 f}{\partial x^j\partial x^i}(x).$$

上面记号中指标的次序指出了函数按怎样的顺序关于相应变量求偏导数.

上面我们定义了二阶偏导数, 如果已经定义了 k 阶偏导数

$$\partial_{i_1\cdots i_k}f(x)=\frac{\partial^k f}{\partial x^{i_1}\cdots\partial x^{i_k}}(x),$$

那么可用归纳法定义 $k+1$ 阶偏导数:

$$\partial_{ii_1\cdots i_k}f(x) := \partial_i(\partial_{i_1\cdots i_k}f)(x).$$

对于多变量函数的高阶偏导数, 发生了一个特殊的问题, 即微分的顺序对要计算的偏导数有影响吗?

定理 3 *设函数 $f:G\to\mathbb{R}$ 在区域 G 上存在偏导数*

$$\frac{\partial^2 f}{\partial x^i\partial x^j}(x),\ \frac{\partial^2 f}{\partial x^j\partial x^i}(x),$$

如果它们在点 $x\in G$ 连续, 则它们在这一点相等.

◀ 设 $x\in G$, 且函数 $\partial_{ij}f:G\to\mathbb{R}$ 与 $\partial_{ji}f:G\to\mathbb{R}$ 在点 x 连续. 下面将在某个球 $B(x;r)\subset G, r>0$(它是点 x 的凸邻域) 上来考察此问题, 我们希望证明:

$$\frac{\partial^2 f}{\partial x^i\partial x^j}(x^1,\cdots,x^m)=\frac{\partial^2 f}{\partial x^j\partial x^i}(x^1,\cdots,x^m).$$

由于在下面的研究中改变的仅仅是变量 x^i 与 x^j, 因此, 为了简单, 可假定 f 是两个变量 x^1 与 x^2 的函数 $f(x^1,x^2)$, 且 $\dfrac{\partial^2 f}{\partial x^1\partial x^2}$ 与 $\dfrac{\partial^2 f}{\partial x^2\partial x^1}$ 在点 $(x^1,x^2)\in B(x;r)\subset G\subset\mathbb{R}^2$ 连续, 要证明:

$$\frac{\partial^2 f}{\partial x^1\partial x^2}(x^1,x^2)=\frac{\partial^2 f}{\partial x^2\partial x^1}(x^1,x^2).$$

考察辅助函数

$$F(h^1,h^2)=f(x^1+h^1,x^2+h^2)-f(x^1+h^1,x^2)-f(x^1,x^2+h^2)+f(x^1,x^2),$$

其中假定增量 $h=(h^1,h^2)$ 足够小, 使得 $x+h\in B(x;r)$.

如果把函数 $F(h^1,h^2)$ 写成差的形式

$$F(h^1,h^2)=\varphi(1)-\varphi(0),$$

其中 $\varphi(t)=f(x^1+th^1,x^2+h^2)-f(x^1+th^1,x^2)$, 那么由拉格朗日定理有

$$F(h^1,h^2)=\varphi'(\theta_1)=(\partial_1 f(x^1+\theta_1h^1,x^2+h^2)-\partial_1 f(x^1+\theta_1h^1,x^2))h^1,$$

对这个差再次应用拉格朗日定理, 有

$$F(h^1,h^2)=\partial_{21}f(x^1+\theta_1h^1,x^2+\theta_2h^2)h^1h^2. \tag{2}$$

如果现在把 $F(h^1,h^2)$ 表成另一差的形式

$$F(h^1,h^2)=\tilde{\varphi}(1)-\tilde{\varphi}(0),$$

其中 $\tilde{\varphi}(t) = f(x^1 + h^1, x^2 + th^2) - f(x^1, x^2 + th^2)$, 那么类似地有

$$F(h^1, h^2) = \partial_{12} f(x^1 + \tilde{\theta}_1 h^1, x^2 + \tilde{\theta}_2 h^2) h^1 h^2. \tag{3}$$

比较等式 (2) 与 (3), 可以断定

$$\partial_{21} f(x^1 + \theta_1 h^1, x^2 + \theta_2 h^2) = \partial_{12} f(x^1 + \tilde{\theta}_1 h^1, x^2 + \tilde{\theta}_2 h^2), \tag{4}$$

其中 $\theta_1, \theta_2, \tilde{\theta}_1, \tilde{\theta}_2 \in]0,1[$. 利用偏导数在点 (x^1, x^2) 的连续性, 在 (4) 式中令 $h \to 0$, 就得到所需要的等式.

$$\partial_{21} f(x^1, x^2) = \partial_{12} f(x^1, x^2) ▶$$

注意, 如果两个偏导数 $\partial_{ij} f$ 与 $\partial_{ji} f$ 在点 x 仅仅存在, 且对它不加任何补充假定, 一般来说, 不能断定等式 $\partial_{ij} f(x) = \partial_{ji} f(x)$ 成立 (参看本节末的习题 2).

我们约定用记号 $C^{(k)}(G;\mathbb{R})$ 或 $C^{(k)}(G)$ 表示定义在区域 $G \subset \mathbb{R}^m$ 上具有直到 k 阶连续偏导数的函数 $f: G \to \mathbb{R}$ 的集合.

作为定理 3 的推论, 有

命题 1 *如果 $f \in C^{(k)}(G;\mathbb{R})$, 则偏导数的值 $\partial_{i_1 \cdots i_k} f(x)$ 不依赖于对 $x^{i_1}, \cdots, x^{i_k}$ 微分的次序, 也就是说, 任意排列指标 $i_1, \cdots, i_k$ 都不会改变偏导数的值.*

◀ 当 $k=2$ 时, 这个命题包含在定理 3 中, 现在假定直到 n 阶命题都成立, 要证明它对 $n+1$ 阶成立.

有 $\partial_{i_1 i_2 \cdots i_{n+1}} f(x) = \partial_{i_1}(\partial_{i_2 \cdots i_{n+1}} f)(x)$, 由归纳假定, 指标 $i_2 \cdots i_{n+1}$ 可以重新排列而不会影响函数 $\partial_{i_2 \cdots i_{n+1}} f$ 在点 x 的值, 从而也不影响函数 $\partial_{i_1 i_2 \cdots i_{n+1}} f$ 在点 x 的值. 因此只要证明, 例如改变 i_1 与 i_2 的次序不会影响导数 $\partial_{i_1 i_2 \cdots i_{n+1}} f(x)$ 就可以了.

因为

$$\partial_{i_1 i_2 \cdots i_{n+1}} f(x) = \partial_{i_1 i_2}(\partial_{i_3 \cdots i_{n+1}} f)(x),$$

那么由定理 3 直接得知 i_1 与 i_2 交换次序不会改变上述导数值. 于是, 根据归纳法, 证明了命题 1. ▶

例 1 设 $f(x) = f(x^1, x^2)$ 是函数类 $C^{(k)}(G;\mathbb{R})$ 中的函数.

设 $h = (h^1, h^2)$, 使得闭区间 $[x, x+h]$ 含在区域 G 中. 我们来证明定义在闭区间 $[0,1] \subset \mathbb{R}$ 上函数

$$\varphi(t) = f(x + th)$$

属于 $C^{(k)}[0,1]$, 并求它对于 t 的 k 阶导数.

我们有

$$\varphi'(t) = \partial_1 f(x^1 + th^1, x^2 + th^2) h^1 + \partial_2 f(x^1 + th^1, x^2 + th^2) h^2,$$

$$\begin{aligned}\varphi''(t) &= \partial_{11} f(x+th) h^1 h^1 + \partial_{21} f(x+th) h^2 h^1 + \partial_{12} f(x+th) h^1 h^2 + \partial_{22} f(x+th) h^2 h^2 \\ &= \partial_{11} f(x+th)(h^1)^2 + 2\partial_{12} f(x+th) h^1 h^2 + \partial_{22} f(x+th)(h^2)^2.\end{aligned}$$

这些关系式可以写成算子 $(h^1\partial_1 + h^2\partial_2)$ 作用于函数的形式:

$$\varphi'(t) = (h^1\partial_1 + h^2\partial_2)f(x+th) = h^i\partial_i f(x+th),$$
$$\varphi''(t) = (h^1\partial_1 + h^2\partial_2)^2 f(x+th) = h^{i_1}h^{i_2}\partial_{i_1 i_2} f(x+th).$$

由归纳法得到

$$\begin{aligned}\varphi^{(k)}(t) &= (h^1\partial_1 + h^2\partial_2)^k f(x+th)\\ &= h^{i_1}\cdot\cdots\cdot h^{i_k}\partial_{i_1\cdots i_k} f(x+th)\end{aligned}$$

(上式右边简记符号指的是关于各种可能的指标 $i_1,\cdots,i_k$ 的和式, 这些指标取数值 1 或 2).

例 2 设 $f(x) = f(x^1,\cdots,x^m)$ 且 $f\in C^{(k)}(G;\mathbb{R})$, 假定 $[x,x+h]\subset G$. 对于定义在闭区间 $[0,1]\subset\mathbb{R}$ 上的函数 $\varphi(t) = f(x+th)$, 我们得到

$$\varphi^{(k)}(t) = h^{i_1}\cdot\cdots\cdot h^{i_k}\partial_{i_1\cdots i_k} f(x+th), \tag{5}$$

其中简记符号指的是关于各种可能的指标 $i_1,\cdots,i_k$ 的和式, 这些指标取从 1 到 m 的所有可能的值.

公式 (5) 也可以写成下面的形式

$$\varphi^{(k)}(t) = (h^1\partial_1 + \cdots + h^m\partial_m)^k f(x+th). \tag{6}$$

4. 泰勒公式

定理 4 *如果函数 $f: U(x)\to\mathbb{R}$ 在点 $x\in\mathbb{R}^m$ 的邻域 $U(x)\subset\mathbb{R}^m$ 上有定义, 并且属于函数类 $C^{(n)}(U(x);\mathbb{R})$, 而 $[x,x+h]\subset U(x)$, 则有*

$$\begin{aligned}&f(x^1+h^1,\cdots,x^m+h^m) - f(x^1,\cdots,x^m)\\ &= \sum_{k=1}^{n-1}\frac{1}{k!}(h^1\partial_1+\cdots+h^m\partial_m)^k f(x) + r_n(x;h),\end{aligned} \tag{7}$$

其中

$$r_{n-1}(x;h) = \int_0^1 \frac{(1-t)^{n-1}}{(n-1)!}(h^1\partial_1+\cdots+h^m\partial_m)^n f(x+th)dt. \tag{8}$$

等式 (7) 与 (8) 一起称为具积分形式余项的泰勒公式.

◀ 泰勒公式 (7) 可立即从单变量函数的泰勒公式得到. 事实上, 考察辅助函数

$$\varphi(t) = f(x+th),$$

由定理 4 的条件知道它定义在闭区间 $0 \leqslant t \leqslant 1$ 上并且属于 $C^{(n)}[0,1]$.

当 $\tau \in [0,1]$ 时, 由单变量函数的泰勒公式可以得到

$$\varphi(\tau) = \varphi(0) + \frac{1}{1!}\varphi'(0)\tau + \cdots + \frac{1}{(n-1)!}\varphi^{(n-1)}(0)\tau^{n-1} + \int_0^1 \frac{(1-t)^{n-1}}{(n-1)!}\varphi^{(n)}(t\tau)dt.$$

令 $\tau = 1$, 有

$$\varphi(1) = \varphi(0) + \frac{1}{1!}\varphi'(0) + \cdots + \frac{1}{(n-1)!}\varphi^{(n-1)}(0) + \int_0^1 \frac{(1-t)^{n-1}}{(n-1)!}\varphi^{(n)}(t)dt. \tag{9}$$

把由公式 (6) 得到的

$$\varphi^{(k)}(0) = (h^1\partial_1 + \cdots + h^m\partial_m)^k f(x) \quad (k = 0, \cdots, n-1),$$
$$\varphi^{(n)}(t) = (h^1\partial_1 + \cdots + h^m\partial_m)^n f(x + th).$$

代入上式就得到定理 4 的结论. ►

注 如果在 (9) 式中用拉格朗日余项代替积分余项, 那么从等式

$$\varphi(1) = \varphi(0) + \frac{1}{1!}\varphi'(0) + \cdots + \frac{1}{(n-1)!}\varphi^{(n-1)}(0) + \frac{1}{n!}\varphi^{(n)}(\theta),$$

其中 $0 < \theta < 1$, 得到带有余项

$$r_{n-1}(x;h) = \frac{1}{n!}(h^1\partial_1 + \cdots + h^m\partial_m)^n f(x + \theta h) \tag{10}$$

的泰勒公式 (7).

这个余项和一个变量函数的情况一样, 称为*具有拉格朗日形式的泰勒公式余项*.

既然 $f \in C^{(n)}(U(x);\mathbb{R})$, 那么从 (10) 式得到

$$r_{n-1}(x;h) = \frac{1}{n!}(h^1\partial_1 + \cdots + h^m\partial_m)^n f(x) + o(\|h\|^n) \quad (h \to 0),$$

因此有

$$f(x^1 + h^1, \cdots, x^m + h^m) - f(x^1, \cdots, x^m) = \sum_{k=0}^{m} \frac{1}{k!}(h^1\partial_1 + \cdots + h^m\partial_m)^k f(x) + o(\|h\|^n) \quad (h \to 0), \tag{11}$$

叫做*具有佩亚诺形式余项的泰勒公式*.

5. 多变量函数的极值 微分学的重要应用之一是利用它寻求和研究函数的极值.

定义 1 设函数 $f: E \to \mathbb{R}$ 定义在集合 $E \subset \mathbb{R}^m$ 上. 如果对于 E 的内点 x_0, 存在邻域 $U(x_0) \subset E$, 当 $x \in U(x_0)$ 时, 有 $f(x) \leqslant f(x_0)(f(x) \geqslant f(x_0))$, 则称函数 f 在 E 的内点 x_0 有*局部极大值* (*局部极小值*).

如果对 $x \in U(x_0)\backslash x_0$ 有严格不等式 $f(x) < f(x_0)(f(x) > f(x_0))$ 则称函数 f 在 x_0 有*严格局部极大值* (*严格局部极小值*).

定义 2 函数的局部极小值与局部极大值称为函数的*局部极值*.

定理 5 *设函数 $f: U(x_0) \to \mathbb{R}$ 定义在点 $x_0 = (x_0^1, \cdots, x_0^m)$ 的邻域 $U(x_0) \subset \mathbb{R}^m$ 上, 并且它在点 x_0 关于每个变量 $x^1, \cdots, x^m$ 存在偏导数. 如果函数 f 在点 x_0 有局部极值, 则有*

$$\frac{\partial f}{\partial x^1}(x_0) = 0, \cdots, \frac{\partial f}{\partial x^m}(x_0) = 0. \tag{12}$$

◀ 考察一个变量的函数 $\varphi(x^1) = f(x^1, x_0^2, \cdots, x_0^m)$, 由定理条件知, 它在实轴上的点 x_0^1 的某个邻域上有定义, 并且它在点 x_0^1 有局部极值, 因此

$$\varphi'(x_0^1) = \frac{\partial f}{\partial x^1}(x_0^1, x_0^2, \cdots, x_0^m) = \frac{\partial f}{\partial x^1}(x_0) = 0.$$

类似地可以证明 (12) 式中其余的等式. ▶

注意, (12) 式仅仅给出多变量函数极值的必要条件, 但不是充分条件, 为了证实这一点, 可以从单变量函数这方面的反例着手. 譬如, 如果以前我们知道函数 $x \mapsto x^3$ 在零点有零导数, 但它在此点无极值, 那么现在就可以考察函数

$$f(x^1, \cdots, x^m) = (x^1)^3.$$

它在点 $x_0 = (0, \cdots, 0)$ 的所有偏导数都等于零, 但是它在此点显然不存在极值.

定理 5 指出, 如果函数 $f: G \to \mathbb{R}$ 定义在开集 $G \subset \mathbb{R}^m$ 上, 那么在它的局部极值点或者 f 不可微, 或者微分 $df(x_0)$ 等于零, 亦即切映射 $f'(x_0)$ 等于零.

我们知道, 如果定义在点 $x_0 \in \mathbb{R}^m$ 的邻域 $U(x_0) \subset \mathbb{R}^m$ 上的映射 $f: U(x_0) \to \mathbb{R}^n$ 在点 x_0 可微, 那么切映射 $f'(x_0): \mathbb{R}^m \to \mathbb{R}^n$ 的矩阵是

$$\begin{pmatrix} \partial_1 f^1(x_0) & \cdots & \partial_m f^1(x_0) \\ \vdots & & \vdots \\ \partial_1 f^n(x_0) & \cdots & \partial_m f^n(x) \end{pmatrix}. \tag{13}$$

定义 3 点 x_0 称为映射 $f: U(x_0) \to \mathbb{R}^n$ 的临界点, 如果映射在这点的雅可比矩阵 (13) 的秩小于 $\min\{m, n\}$, 也就是说, 比矩阵 (13) 可能的最大秩要小.

特别地，当 $n=1$ 时，如果满足条件 (12)，亦即函数 $f: U(x_0) \to \mathbb{R}$ 的所有偏导数都等于零，则点 x_0 为临界点.

实值函数的临界点同时也称为这个函数的稳定点.

我们通过解方程组 (12) 求得函数的临界点，然后进一步分析它们是不是函数的极值点. 这个工作得以顺利进行是利用由泰勒公式以及由它得到的下面的极值是否存在的充分条件.

定理 6 设 $f: U(x_0) \to \mathbb{R}$ 是定义在点 $x_0=(x_0^1,\cdots,x_0^m)$ 的邻域 $U(x_0)\subset\mathbb{R}^m$ 上且属于 $C^{(2)}(U(x_0);\mathbb{R})$ 的函数，点 x_0 是函数 f 的临界点. 如果函数在 x_0 的泰勒展开式

$$f(x_0^1+h^1,\cdots,x_0^m+h^m)=f(x_0^1,\cdots,x_0^m)+\frac{1}{2!}\sum_{i,j=1}^{m}\frac{\partial^2 f}{\partial x^i\partial x^j}(x_0)h^ih^j+o(\|h\|^2) \quad (14)$$

中的二次型

$$\sum_{i,j=1}^{m}\frac{\partial^2 f}{\partial x^i\partial x^j}(x_0)h^ih^j\equiv\partial_{ij}f(x_0)h^ih^j \quad (15)$$

a) 有确定的符号，则函数 f 在 x_0 取得局部极值. 当二次型 (15) 是正定的时候，f 在 x_0 取得严格局部极小值；当二次型 (15) 是负定的时候，f 在 x_0 取得严格局部极大值.

b) 具有不同的符号，则函数 f 在 x_0 没有极值.

◀ 设 $h\neq 0, x_0+h\in U(x_0)$，把 (14) 式写成如下形式

$$f(x_0+h)-f(x_0)=\frac{1}{2}\|h\|^2\left[\sum_{i,j=1}^{m}\frac{\partial^2 f}{\partial x^i\partial x^j}(x_0)\frac{h^i}{\|h\|}\frac{h^j}{\|h\|}+o(1)\right], \quad (16)$$

其中 $o(1)$ 当 $h\to 0$ 时是无穷小量.

从 (16) 式看出，差式 $f(x_0+h)-f(x_0)$ 的符号完全由右端方括号内的量的符号决定，我们现在来研究它.

向量 $e=\left(\frac{h^1}{\|h\|},\cdots,\frac{h^m}{\|h\|}\right)$ 的范数显然是 1，二次型 (15) 看成 h 的函数，在 $\mathbb{R}^m$ 上连续，因此它在单位球面

$$S(0;1)=\{x\in\mathbb{R}^m \mid \|x\|=1\}$$

上连续且有界，但是球面 S 是 $\mathbb{R}^m$ 中的有界闭集，即紧集，于是二次型 (15) 在 S 上取得最小值 m 与最大值 M.

如果二次型 (15) 是正定的，那么有 $0<m\leqslant M$，并且可找到一个 $\delta>0$，使得当 $\|h\|<\delta$ 时有 $|o(1)|<m$. 这时 (16) 式右边方括号中的量保持正号，也就是说，当

$0<\|h\|<\delta$ 时, 有 $f(x_0+h)-f(x_0)>0$, 因此, 点 x_0 是函数 f 的严格局部极小值点.

类似地可以证明, 当二次型 (15) 是负定时, 函数 f 在 x_0 取得严格局部极大值.

这样, 定理 6 的 a) 就证完了, 现在来证明结论 b).

设 e_m, e_M 是单位球面上的点, 二次型 (15) 在 e_m, e_M 分别取值 m, M, 而且 $m<0<M$. 令 $h=te_m$, 其中 $t>0$ 充分小, 使得 $x_0+h\in U(x_0)$, 从 (16) 式得到

$$f(x_0+te_m)-f(x_0)=\frac{1}{2!}t^2(m+o(1)),$$

其中 $o(1)\to 0$ 当 $t\to 0$. 而且, 该等式右端的 $m+o(1)$, 当 t 充分小时, 与 m 的符号相同, 即负号, 因而上式左边也取负号.

类似地令 $h=te_M$, 得

$$f(x_0+te_M)-f(x_0)=\frac{1}{2!}t^2(M+o(1)),$$

从而, 当 t 充分小时, 差式 $f(x_0+te_M)-f(x_0)$ 是正的.

这样, 如果二次型 (15) 在单位球面上, 等价地, 在 $\mathbb{R}^m$ 上具有不同的符号, 那么在点 x_0 的任一邻域内可以找到一个点, 使得函数在此点的值大于 $f(x_0)$, 同时又可找到另一个点, 使得函数在此点之值小于 $f(x_0)$. 因此, 点 x_0 不是函数 f 的局部极值点. ►

现在叙述与定理有关的若干附注.

注 1 定理 6 没有讨论二次型 (15) 是半定的情形, 即非正且或非负的情形, 在这种情形, 点 x_0 可以是函数的极值点也可以不是, 这可从下面例子看出来.

例 3 求定义在 $\mathbb{R}^2$ 上的函数 $f(x,y)=x^4+y^4-2x^2$ 的极值.

按照必要条件 (12) 写出方程组

$$\begin{cases}\dfrac{\partial f}{\partial x}(x,y)=4x^3-4x=0,\\ \dfrac{\partial f}{\partial y}(x,y)=4y^3=0.\end{cases}$$

求得三个临界点:$(-1,0),(0,0),(1,0)$.

因为

$$\frac{\partial^2 f}{\partial x^2}(x,y)=12x^2-4,\ \frac{\partial^2 f}{\partial x\partial y}(x,y)\equiv 0,\ \frac{\partial^2 f}{\partial y^2}(x,y)=12y^2,$$

那么二次型 (15) 在上面三个临界点处分别有值

$$8(h^1)^2, -4(h^1)^2, 8(h^1)^2,$$

也就是说, 在所有三种情况它都是半定的 (非正的或非负的), 这时不能应用定理 6, 但是因为 $f(x,y)=(x^2-1)^2+y^4-1$, 显然看出, 函数 f 在点 $(-1,0)$ 与 $(1,0)$ 处有严

格极小值 -1(甚至不是局部的), 而在点 $(0,0)$ 它没有极值, 这是因为当 $x=0, y\neq 0$ 时, $f(0,y)=y^4>0$, 而当 $y=0, x$ 充分小但不为零时, $f(x,0)=x^4-2x^2<0$.

注 2 当二次型 (15) 得到以后, 可以利用代数教材中有名的西尔维斯特准则[①]研究它的确定性. 根据这个准则, 具有对称矩阵

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix}$$

的二次型 $\sum\limits_{i,j=1}^{m} a_{ij}x^ix^j$ 正定, 当且仅当, 它的矩阵的所有主子式都大于零; 它负定, 当且仅当, $a_{11}<0$, 且每一主子式与下一阶主子式有相反的符号.

例 4 求在 $\mathbb{R}^2$ 上, 除去坐标原点, 处处有定义的函数

$$f(x,y)=xy\ln(x^2+y^2)$$

的极值.

解方程组

$$\begin{cases} \dfrac{\partial f}{\partial x}(x,y)=y\ln(x^2+y^2)+\dfrac{2x^2y}{x^2+y^2}=0, \\ \dfrac{\partial f}{\partial y}(x,y)=x\ln(x^2+y^2)+\dfrac{2xy^2}{x^2+y^2}=0, \end{cases}$$

得到函数所有的临界点:

$$(0,\pm1),(\pm1,0),\left(\pm\frac{1}{\sqrt{2e}},\pm\frac{1}{\sqrt{2e}}\right),\left(\pm\frac{1}{\sqrt{2e}},\mp\frac{1}{\sqrt{2e}}\right).$$

因为函数关于两个自变量中的每一个都是奇函数, 所以点 $(0,\pm1)$ 与 $(\pm1,0)$ 显然不是函数的极值点. 另外, 当两个变量 x 与 y 同时改变符号时, 函数值不会改变, 因此, 我们只要再研究剩下来的临界点中的一个, 例如点 $\left(\frac{1}{\sqrt{2e}},\frac{1}{\sqrt{2e}}\right)$, 就可以推断其他点的性质了.

因为

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(x,y)&=\frac{6xy}{x^2+y^2}-\frac{4x^3y}{(x^2+y^2)^2}, \\ \frac{\partial^2 f}{\partial x\partial y}(x,y)&=\ln(x^2+y^2)+2-\frac{4x^2y^2}{(x^2+y^2)^2}, \\ \frac{\partial^2 f}{\partial y^2}(x,y)&=\frac{6xy}{x^2+y^2}-\frac{4xy^3}{(x^2+y^2)^2}. \end{aligned}$$

① 西尔维斯特 (J. J. Sylvester)(1814—1897) —— 英国数学家, 他最有名的工作是在代数方面.

所以二次型 $\partial_{ij}f(x_0)h^ih^j$ 在点 $\left(\frac{1}{\sqrt{2e}}, \frac{1}{\sqrt{2e}}\right)$ 的矩阵为

$$\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

因此, 它是正定的, 从而, 函数在此点有局部极小值 $f\left(\frac{1}{\sqrt{2e}}, \frac{1}{\sqrt{2e}}\right) = -\frac{1}{2e}$.

由上面看到的所考察的函数的特点, 可以断定

$$f\left(\frac{-1}{\sqrt{2e}}, \frac{-1}{\sqrt{2e}}\right) = \frac{-1}{2e}$$

也是函数的极小值, 而

$$f\left(\frac{1}{\sqrt{2e}}, \frac{-1}{\sqrt{2e}}\right) = f\left(\frac{-1}{\sqrt{2e}}, \frac{1}{\sqrt{2e}}\right) = \frac{1}{2e}$$

是函数的局部极大值. 其实, 这些结论都可以直接由二次型的确定性来验证. 例如, 二次型 (15) 在点 $\left(\frac{-1}{\sqrt{2e}}, \frac{-1}{\sqrt{2e}}\right)$ 的矩阵为

$$\begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$

由此看出二次型是负定的.

注 3 应当注意, 我们得到的关于函数极值的必要条件 (定理 5) 与充分条件 (定理 6) 仅仅适用于函数定义域中的内点. 这样, 在寻求函数的最大值或最小值时, 除了必须研究内部的临界点外, 还要研究定义域的边界点, 因为函数的最大值或最小值可能在边界点取到.

更详细地研究函数在非内点处极值的一般法则将在以后叙述 (参看条件极值一节). 在寻求函数的极大值与极小值时, 除正式的作法外, 有时可利用与问题性质有关的一些简单的想法. 注意到这一点是有益的. 例如, 如果所考察的在 $\mathbb{R}^m$ 上可微的函数, 按问题的意义应当有极小值, 它在 $\mathbb{R}^m$ 上无上界, 且有唯一的临界点, 那么可以不作进一步研究就能断定函数在此点具有极小值.

例 5 惠更斯问题 根据在封闭力学系统中的能量和动量守恒原理, 经过不太复杂的计算可以证明, 两个质量为 m_1 与 m_2, 初速度为 v_1 与 v_2 的绝对弹性的球发生对心完全弹性碰撞以后, 它们的速度 (速度方向沿着两球中心的连线) 由下面的式子确定:

$$\tilde{v}_1 = \frac{(m_1 - m_2)v_1 + 2m_2v_2}{m_1 + m_2},$$
$$\tilde{v}_2 = \frac{(m_2 - m_1)v_2 + 2m_1v_1}{m_1 + m_2}.$$

特别地, 如果质量为 M, 速度为 V 的球与一个质量为 m 的不动的球碰撞, 那么从上面公式中可以求出质量为 m 的小球获得的速度为

$$v = \frac{2M}{m+M}V, \tag{17}$$

由此看出, 若 $0 \leqslant m \leqslant M$, 则 $V \leqslant v \leqslant 2V$. 用什么方法可以把大质量物体的较多的动能转移到质量小的物体呢? 例如, 可以在小质量球与大质量球之间放入具有中间质量 $m < m_1 < m_2 < \cdots < m_n < M$ 的球, 令具大质量 M 的球与质量为 $m_n, m_{n-1}, \cdots, m_2, m_1, m$ 的球依次对心完全弹性碰撞, 如何选取 $m_1, m_2, \cdots, m_n$ 才能使具小质量 m 的球获得最大的速度呢 (由惠更斯提出)?

按照公式 (17), 小球获得的速度 v 可表为变量 $m_1, m_2, \cdots, m_n$ 的函数:

$$v = \frac{m_1}{m+m_1} \cdot \frac{m_2}{m_1+m_2} \cdot \cdots \cdot \frac{m_n}{m_{n-1}+m_n} \cdot \frac{M}{m_n+M} \cdot 2^{n+1}V, \tag{18}$$

这样, 惠更斯问题归结为寻求函数

$$f(m_1, \cdots, m_n) = \frac{m_1}{m+m_1} \cdot \frac{m_2}{m_1+m_2} \cdot \cdots \cdot \frac{m_n}{m_{n-1}+m_n} \cdot \frac{M}{m_n+M}$$

的极大值.

内部极值必要条件方程组 (12) 在本题中归结为

$$\begin{cases} m \cdot m_2 - m_1^2 = 0, \\ m_1 \cdot m_3 - m_2^2 = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ m_{n-1} \cdot M - m_n^2 = 0. \end{cases}$$

由此看出, 数

$$m, m_1, m_2, \cdots, m_n, M$$

构成等比数列, 公比 $q = \sqrt[n+1]{\dfrac{M}{m}}$.

把这样得到的质量组代入 (18), 得到

$$v = \left(\frac{2q}{1+q}\right)^{n+1} V, \tag{19}$$

当 $n = 0$ 时, 上式与 (17) 式重合.

从物理意义考虑, 公式 (19) 显然给出函数 (18) 的最大值, 但是这也可以正式地予以证明 (并不需要进行繁琐的二阶导数的计算, 参看本节末的习题 9).

注意, 从公式 (19) 看出, 如果 $m \to 0$, 则 $v \to 2^{n+1}V$, 就是说, 放在小质量 m 与大质量 M 之间的一系列球确实明显地增加了大质量 M 的动能向小质量 m 的物体

的转移. ①

6. 与多变量函数有关的某些几何形象

a. 函数图像与曲线坐标 设 x, y, z 是空间 $\mathbb{R}^3$ 中点的笛卡儿坐标, $z = f(x, y)$ 是定义在平面 $\mathbb{R}^2$ 的某个区域 G 上的变量 (x, y) 的连续函数.

由函数图像的一般定义知道, 这时函数 $f : G \to \mathbb{R}$ 的图像是空间 $\mathbb{R}^3$ 中的集合

$$S = \{(x, y, z) \in \mathbb{R}^3 | (x, y) \in G, z = f(x, y)\}.$$

用关系式 $(x, y) \mapsto (x, y, f(x, y))$ 定义的映射 $G \xrightarrow{F} S$ 显然是连续的相互单值的从 G 到 S 上的映射, 根据它, 集合 S 上任一点 $(x, y, f(x, y))$ 都可通过指出与它对应的区域 G 的点或这个点的坐标 (x, y) 的方法指定.

这样一来, 可以把数对 (x, y) 看成是集合 S(即函数 $z = f(x, y)$ 的图像) 上点的一种坐标. 因为 S 的点是用数对 (x, y) 给出的, 故称 S 为 $\mathbb{R}^3$ 中的二维曲面(曲面的一般定义以后再给出.)

如果给出 G 中的道路 $\Gamma : I \to G$, 那么自然得到曲面S 上的道路 $F \circ \Gamma : I \to S$, 如果道路 Γ 用参数式 $x = x(t), y = y(t)$ 给出, 那么 S 上的道路 $F \circ \Gamma$ 用三个函数 $x = x(t), y = y(t), z = f(x(t), y(t))$ 给出. 特别地, 如果设 $x = x_0 + t, y = y_0$, 那么我们得到曲面 S 上的曲线 $x = x_0 + t, y = y_0, z = f(x_0 + t, y_0)$, 沿着这条曲线, 曲

---

①译者注: 为了准确了解动能转移与中间球的个数的关系, 可作如下分析.

将 $q = \sqrt[n+1]{M/m}$ 代入 (19) 式, 得

$$v = \left[\frac{2}{1 + \left(\frac{m}{M}\right)^{1/(n+1)}}\right]^{n+1} \cdot V.$$

注意到 $M > m$, 容易验证, 它关于 n 是单调增加的, 就是说, 在大、小球之间设置的球越多, 则从大球转移到小球的动能就越多, 利用洛必达法则, 得

$$\begin{aligned}
&\lim_{n\to\infty} \ln\left[\frac{2}{1 + \left(\frac{m}{M}\right)^{\frac{1}{n+1}}}\right]^{n+1} \\
&= \lim_{n\to\infty} (n+1)\left\{\ln 2 - \ln\left[1 + \left(\frac{m}{M}\right)^{\frac{1}{n+1}}\right]\right\} \\
&= \lim_{x\to 0} \frac{\ln 2 - \ln\left[1 + \left(\frac{m}{M}\right)^x\right]}{x} \\
&= \lim_{x\to 0} \frac{-\left(\frac{m}{M}\right)^x / \ln\frac{m}{M}}{1 + \left(\frac{m}{M}\right)^x} = \ln\sqrt{\frac{M}{m}}.
\end{aligned}$$

因此, $\lim\limits_{n\to\infty} \frac{1}{2}mv^2 - \frac{1}{2}MV^2$. 这就是说, 当 $n \quad \to \infty$, 大球转移给小球的动能将趋于大球原来的全部动能

面 S 上点 (x,y,z) 的坐标 $y=y_0$ 保持不变. 类似地, 可以得到曲面 S 上另一条曲线 $x=x_0, y=y_0+t, z=f(x_0,y_0+t)$, 沿着它, 曲面 S 上点 (x,y,z) 的第一个坐标 $x=x_0$ 保持不变. 曲面 S 上的这些线, 类似于平面情况, 自然称为曲面 S 上的坐标线. $G\subset\mathbb{R}^2$ 中的坐标线是一段直线, 曲面 S 上的坐标线不同, 一般说是 $\mathbb{R}^3$ 中的曲线. 由于这个原因, 曲面 S 上点的坐标 (x,y) 常称为 S 上的曲线坐标.

因此, 定义在区域 $G\subset\mathbb{R}^2$ 上的连续函数 $z=f(x,y)$ 的图像是 $\mathbb{R}^3$ 中的二维曲面, 它的点可以用曲线坐标 $(x,y)\in G$ 表示.

我们暂时不讨论曲面的一般定义, 因为现在我们的兴趣仅仅是曲面的特殊情况——函数的图像, 但是, 我们假定读者从解析几何教程中已经很熟悉 $\mathbb{R}^3$ 中某些重要的具体的曲面 (平面、椭球面、抛物面、双曲面).

b. 函数图像的切平面 如果函数 $z=f(x,y)$ 在点$(x_0,y_0)\in G$ 可微, 那么有

$$\begin{aligned}f(x,y)=&f(x_0,y_0)+A(x-x_0)+B(y-y_0)\\&+o\left(\sqrt{(x-x_0)^2+(y-y_0)^2}\,\right)\quad((x,y)\to(x_0,y_0)).\end{aligned}\tag{20}$$

其中 A,B 是常数. 在 $\mathbb{R}^3$ 中考察平面

$$z=z_0+A(x-x_0)+B(y-y_0),\tag{21}$$

其中 $z_0=f(x_0,y_0)$. 比较等式 (20) 与 (21) 后看出, 我们的函数图像在点 (x_0,y_0,z_0) 的邻域内非常接近于平面 (21). 确切些说, 函数图像上的点 $(x,y,f(x,y))$ 与平面 (21) 上的点 $(x,y,z(x,y))$ 的偏差与量 $\sqrt{(x-x_0)^2+(y-y_0)^2}$(它是从点 (x_0,y_0,z_0) 的坐标 (x_0,y_0) 到图像上点 $(x,y,f(x,y))$ 的曲线坐标 (x,y) 的位移) 比较是一个无穷小量.

由函数微分的唯一性知道, 具有上述性质的平面 (21) 是唯一的并有如下形式

$$z=f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0)+\frac{\partial f}{\partial y}(x_0,y_0)(y-y_0).\tag{22}$$

它称为函数 $z=f(x,y)$ 图像在点 $(x_0,y_0,f(x_0,y_0))$ 的切平面.

于是, 函数 $z=f(x,y)$ 在点 (x_0,y_0) 的可微性与这个函数的图像在点 $(x_0,y_0,f(x_0,y_0))$ 存在切平面是等价的.

c. 法向量 把切平面方程 (22) 写成标准形式

$$\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0)+\frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)-(z-f(x_0,y_0))=0,$$

由此可知, 向量

$$\left(\frac{\partial f}{\partial x}(x_0,y_0),\frac{\partial f}{\partial y}(x_0,y_0),-1\right)\tag{23}$$

是切平面的法向量. 它的方向看作是曲面S (函数图像) 在点 $(x_0, y_0, f(x_0, y_0))$ 的法向或正交于 S 的方向.

特别地, 如果点 (x_0, y_0) 是函数 $f(x, y)$ 的临界点, 那么图像在点 $(x_0, y_0, f(x_0, y_0))$ 处的法向量为 $(0, 0, -1)$, 从而, 函数图像在这个点的切平面是水平的 (平行于平面 (x, y)).

下面的三个函数图像可以说明上边所说的内容.

图 53 中的 (a) 与 (c) 描绘了在局部极值点邻域中函数的图像相对于它的切平面的位置. 在图 53 中的 (b) 画出了函数图像在所谓鞍形临界点邻域的情形

d. 切平面和切向量 我们知道, 如果 $\mathbb{R}^3$ 中的道路 $\Gamma : I \to \mathbb{R}^3$ 的表示式 $x = x(t), y = y(t), z = z(t)$ 是可微函数, 那么向量 $(\dot{x}(0), \dot{y}(0), \dot{z}(0))$ 是时刻 $t = 0$ 的速度向量. 这个向量通过道路 Γ 的承载子上的点 $x_0 = x(0), y_0 = y(0), z_0 = z(0)$ 且与曲线相切.

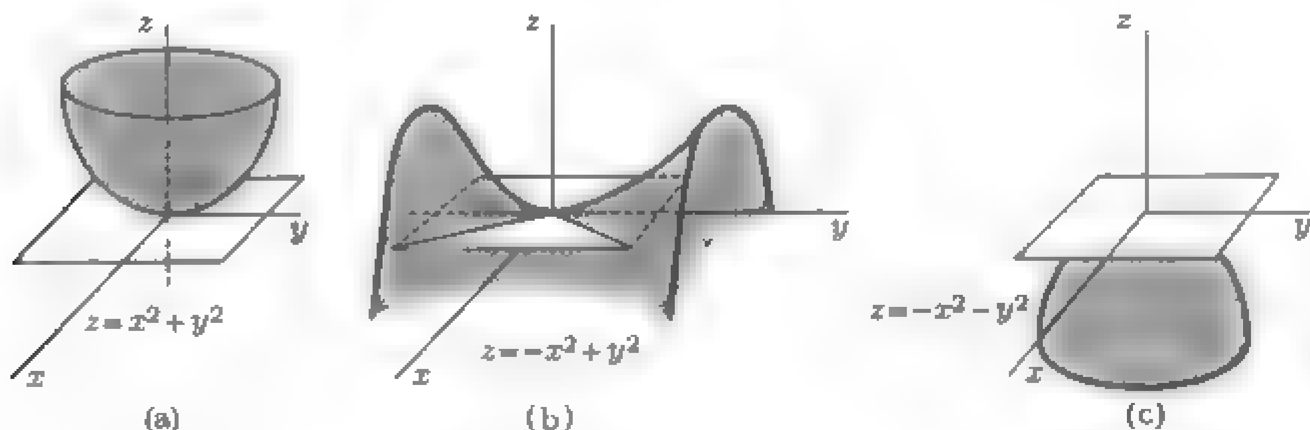


图 53

现在考察函数 $z = f(x, y)$ 图像上的道路 $\Gamma : I \to S$, 它的表示式为 $x = x(t), y = y(t), z = f(x(t), y(t))$. 这时有

$$(\dot{x}(0), \dot{y}(0), \dot{z}(0)) = \left(\dot{x}(0), \dot{y}(0), \frac{\partial f}{\partial x}(x_0, y_0)\dot{x}(0) + \frac{\partial f}{\partial y}(x_0, y_0)\dot{y}(0)\right),$$

由此看出, 上述向量垂直于向量 (23), 即垂直于函数图像在点 $(x_0, y_0, f(x_0, y_0))$ 的法向量. 这样, 我们证明了, 如果向量 (ξ, η, ζ) 在点 $(x_0, y_0, f(x_0, y_0))$ 与曲面S 上过此点的某曲线相切, 那么它垂直于向量 (23) 并且 (在这种意义下) 在平面 (22) 上, 这个平面在所指出的点与曲面 S 相切. 更确切地可以说, 所有直线

$$x = x_0 + \xi t, \quad y = y_0 + \eta t, \quad z = f(x_0, y_0) + \zeta t$$

都在切平面 (22) 上.

现在我们要证明, 相反的命题也是对的, 也就是说, 如果直线 $x = x_0 + \xi t, y = y_0 + \eta t, z = f(x_0, y_0) + \zeta t$, 或向量 (ξ, η, ζ) 位于函数 $z = f(x, y)$ 图像 S 在点

$(x_0, y_0, f(x_0, y_0))$ 的切平面 (22) 上, 那么在 S 上存在一条道路, 使得向量 (ξ, η, ζ) 是它在点 $(x_0, y_0, f(x_0, y_0))$ 的速度向量.

作为这样的道路, 可以取

$$x = x_0 + \xi t, y = y_0 + \eta t, \quad z = f(x_0 + \xi t, y_0 + \eta t).$$

事实上, 对于它, 有

$$\dot{x}(0) = \xi, \dot{y}(0) = \eta, \dot{z}(0) = \frac{\partial f}{\partial x}(x_0, y_0)\xi + \frac{\partial f}{\partial y}(x_0, y_0)\eta,$$

由于

$$\frac{\partial f}{\partial x}(x_0, y_0)\dot{x}(0) + \frac{\partial f}{\partial y}(x_0, y_0)\dot{y}(0) - \dot{z}(0) = 0,$$

且按条件还有

$$\frac{\partial f}{\partial x}(x_0, y_0)\xi + \frac{\partial f}{\partial y}(x_0, y_0)\eta - \zeta = 0,$$

我们断定

$$(\dot{x}(0), \dot{y}(0), \dot{z}(0)) = (\xi, \eta, \zeta).$$

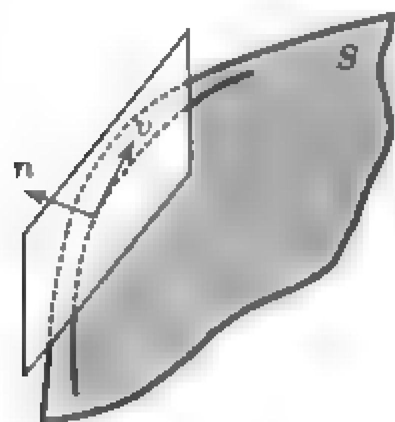


图 54

于是, 曲面S 在点 (x_0, y_0, z_0) 处的切平面是由通过 S 上的点 (x_0, y_0, z_0) 且与曲面上过这个点的曲线相切的向量组成的 (图 54).

这已是切平面的比较几何化的描述了. 无论如何, 从它可以看出, 如果曲线的切线定义关于坐标系的取法保持不变, 那么切平面定义同样也保持不变.

为了直观, 我们仅考察了两个变量的函数, 然而以上的一切叙述显然可以推广到一般的 $m \in \mathbb{N}$ 个变量的函数

$$y = f(x^1, \cdots, x^m). \tag{24}$$

这个函数的图像在点 $(x_0^1, \cdots, x_0^m, f(x_0^1, \cdots, x_0^m))$ 处的切平面是

$$y = f(x_0^1, \cdots, x_0^m) + \sum_{i=1}^{m} \frac{\partial f}{\partial x^i}(x_0^1, \cdots, x_0^m)(x^i - x_0^i). \tag{25}$$

向量

$$\left(\frac{\partial f}{\partial x^1}(x_0), \cdots, \frac{\partial f}{\partial x^m}(x_0), -1\right)$$

是平面 (25) 的法向量. 这个平面本身如同函数 (24) 的图像一样有相同的维数 m, 即它们的任一点都由 m 个坐标的有序组 $(x^1, \cdots, x^m)$ 决定.

这样, 方程 (25) 给出了 $\mathbb{R}^{m+1}$ 中的超平面.

逐字逐句重复上面的推导, 可以断定, 切平面 (25) 是由 m 维曲面 S 即函数 (24) 的图像上通过点 $(x_0^1, \cdots, x_0^m, f(x_0^1, \cdots, x_0^m))$ 的曲线的切向量组成的.

练　习

1. 设 $z=f(x,y)$ 是定义在区域 $G\subset\mathbb{R}^2$ 上的连续可微函数类 $C^{(1)}(G;\mathbb{R})$ 中的函数.

a) 如果 $\dfrac{\partial f}{\partial y}(x,y)=0,(x,y)\in G$, 那么可以断定函数 f 在 G 上不依赖于 y 吗?

b) 区域 G 具备什么条件, 才能使得上述问题有肯定的回答?

2. a) 证明: 设函数

$$f(x,y)=\begin{cases}xy\dfrac{x^2-y^2}{x^2+y^2}&,x^2+y^2\neq 0,\\ 0&,x^2+y^2=0.\end{cases}$$

则有

$$\frac{\partial^2 f}{\partial x\partial y}(0,0)=1\neq -1=\frac{\partial^2 f}{\partial y\partial x}(0,0).$$

b) 证明: 如果函数 $f(x,y)$ 在点 (x_0,y_0) 的邻域 U 上有偏导数 $\dfrac{\partial f}{\partial x},\dfrac{\partial f}{\partial y}$, 混合导数 $\dfrac{\partial^2 f}{\partial x\partial y}$ $\left(\text{或}\dfrac{\partial^2 f}{\partial y\partial x}\right)$ 在 U 上存在且在点 (x_0,y_0) 连续, 则混合导数 $\dfrac{\partial^2 f}{\partial y\partial x}$ $\left(\text{相应地}\dfrac{\partial^2 f}{\partial x\partial y}\right)$ 在这点也存在且有

$$\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0)=\frac{\partial^2 f}{\partial y\partial x}(x_0,y_0).$$

3. 设 $x^1,\cdots,x^m$ 是 $\mathbb{R}^m$ 中的笛卡儿坐标, $f\in C^{(2)}(G;\mathbb{R})$. 用式子

$$\Delta f=\sum_{i=1}^{m}\frac{\partial^2 f}{\partial x^{i^2}}(x^1,\cdots,x^m)$$

定义作用在 $C^{(2)}(G;\mathbb{R})$ 上的微分算子

$$\Delta=\sum_{i=1}^{m}\frac{\partial^2}{\partial x^{i^2}},$$

称其为*拉普拉斯算子*.

在区域 $G\subset\mathbb{R}^m$ 中关于函数 f 的方程 $\Delta f=0$ 称为*拉普拉斯方程*，它的解称为区域 G 中的*调和函数*.

a) 证明: 设 $x=(x^1,\cdots,x^m),\|x\|=\sqrt{\sum\limits_{i=1}^{m}(x^i)^2}$, 则当 $m>2$ 时, 函数

$$f(x)=\|x\|^{-\frac{2-m}{2}}$$

在区域 $\mathbb{R}^m\setminus 0$ (其中 $0=(0,\cdots,0)$) 中是调和的.

b) 证明: 函数

$$f(x^1,\cdots,x^m,t)=\frac{1}{(2a\sqrt{\pi t})^m}e^{-\frac{\|x\|^2}{4a^2t}}$$

当 $t>0$ 且 $x=(x^1,\cdots,x^m)\in\mathbb{R}^m$ 时有定义且满足*热传导方程*

$$\frac{\partial f}{\partial t}=a^2\Delta f,$$

即在函数定义域中任一点有 $\dfrac{\partial f}{\partial t}=a^2\sum\limits_{i=1}^{m}\dfrac{\partial^2 f}{\partial x^{i^2}}$.

4. 多重指标记号的泰勒公式 记 $\alpha := (\alpha_1, \cdots, \alpha_m)$, 它由非负整数 $\alpha_i (i = 1, \cdots, m)$ 组成, 称为多重指标α. 又记

$$|\alpha| := |\alpha_1| + \cdots + |\alpha_m|,$$
$$\alpha! := \alpha_1! \cdot \cdots \cdot \alpha_m!;$$

最后, 若 $a = (a_1, \cdots, a_m)$, 则记

$$a^\alpha := a_1^{\alpha_1} \cdot \cdots \cdot a_m^{\alpha_m}.$$

a) 证明: 若 $k \in \mathbb{N}$, 则

$$(a_1 + \cdots + a_m)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha_1! \cdot \cdots \cdot \alpha_m!} a_1^{\alpha_1} \cdot \cdots \cdot a_m^{\alpha_m},$$

或

$$(a_1 + \cdots + a_m)^k = \sum_{|\alpha|=k} \frac{k!}{\alpha!} a^\alpha,$$

这里, 求和号是对所有满足等式 $\sum_{i=1}^{m} |\alpha_i| = k$ 的非负数组 $\alpha = (\alpha_1, \cdots, \alpha_m)$ 取的.

b) 设

$$D^\alpha f(x) := \frac{\partial^{|\alpha|} f}{(\partial x^1)^{\alpha_1} \cdot \cdots \cdot (\partial x^m)^{\alpha_m}}(x).$$

证明: 若 $f \in C^{(k)}(G; \mathbb{R})$, 则对任意的点 $x \in G$, 有等式

$$\sum_{i_1 + \cdots + i_m = k} \partial_{i_1 \cdots i_k} f(x) h^{i_1} \cdot \cdots \cdot h^{i_k} = \sum_{|\alpha|=k} \frac{k!}{\alpha!} D^\alpha f(x) h^\alpha,$$

其中 $h = (h^1, \cdots, h^m)$.

c) 证明: 带有拉格朗日余项的多重指标记号的泰勒公式为

$$f(x+h) = \sum_{|\alpha|=0}^{n-1} \frac{1}{\alpha!} D^\alpha f(x) h^\alpha + \sum_{|\alpha|=n} \frac{1}{\alpha!} D^\alpha f(x + \theta h) h^\alpha.$$

d) 用多重指标记号写出带积分余项的泰勒公式 (定理 4).

5. a) 设 $I^m = \{x = (x^1, \cdots, x^m) \in \mathbb{R}^m \mid |x^i| \leqslant c^i, i = 1, \cdots, m\}$ 是 m 维区间, I 是闭区间 $[a, b] \subset \mathbb{R}$. 证明: 如果函数 $f(x, y) = f(x^1, \cdots, x^m, y)$ 在集合 $I^m \times I$ 上有定义且连续, 则对任意的 $\varepsilon > 0$, 存在 $\delta > 0$, 使得当 $x \in I^m, y_1, y_2 \in I$ 且 $|y_1 - y_2| < \delta$, 恒有

$$|f(x, y_1) - f(x, y_2)| < \varepsilon.$$

b) 证明: 函数

$$F(x) = \int_a^b f(x, y) dy$$

在区间 I^m 上定义且连续.

c) 证明: 若 $f \in C(I^m;\mathbb{R})$, 则函数

$$\mathcal{F}(x,t) = f(tx)$$

在 $I^m \times I^1$ 上定义且连续, 其中 $I^1 = \{t \in \mathbb{R} \mid |t| \leqslant 1\}$.

d) 证明阿达马引理: 若 $f \in C^{(1)}(I^m;\mathbb{R})$ 且 $f(0) = 0$, 则存在函数 $g_1, \cdots, g_m \in C(I^m;\mathbb{R})$, 使得在 I^m 中成立

$$f(x^1, \cdots, x^m) = \sum_{i=1}^{m} x^i g_i(x^1, \cdots, x^m),$$

且

$$g_i(0) = \frac{\partial f}{\partial x^i}(0), (i = 1, \cdots, m).$$

6. 证明下面的罗尔定理对于多变量函数的推广: *如果函数 f 在闭球 $\bar{B}(0;r)$ 上连续, 在它的边界上等于零并且在球 $B(0;r)$ 的内点可微, 则这个球至少有一个内点是函数的临界点.*

7. 证明: 函数

$$f(x,y) = (y - x^2)(y - 3x^2)$$

在坐标原点没有极值, 虽然它在任何一条通过坐标原点的直线上的限制在原点取得局部严格极小值.

8. **最小二乘法** 这是广泛采用的处理观察数据的一种方法. 设已知物理量 x 与 y 之间存在线性关系

$$y = ax + b \tag{26}$$

要求根据实验结果来构造这个形式的经验公式.

假定做了 n 次观察, 每次测量同时得到两个值 x 和 y, 由此得到下面一列成对的值: x_1, y_1; $\cdots$; x_n, y_n, 由于测量有误差, 因此, 即使两个量 x 和 y 之间有精确的关系式 (26), 可无论系数 a 和 b 是怎样的数, 等式

$$y_k = ax_k + b$$

($k \in \{1, \cdots, n\}$) 中仍可能有不成立的.

问题是如何根据观测结果用合理的方式确定未知系数 a 和 b.

高斯根据对观测误差的概率分布的分析, 建立了在给定全部观察结果的情况下, 寻求系数 a 与 b 的最大概率值的方法, 方法基于下面的**最小二乘法原理**:

如果 $\delta_k = (ax_k + b) - y_k$ 是第 k 次观察的闭合差, 那么 a 与 b 必须选取得使闭合差的平方和

$$\Delta = \sum_{k=1}^{n} \delta_k^2$$

最小.

a) 证明: 在关系 (26) 的情形, 最小二乘法原理导致下述关于系数 a, b 的线性方程组

$$a[x_k, x_k] + b[1, x_k] = [x_k, y_k],$$
$$a[x_k, 1] + b[1, 1] = [y_k, 1],$$

这里, 遵循高斯, 记 $[x_k, x_k] := x_1x_1+\cdots+x_nx_n$, $[1, x_k] := 1\cdot x_1+\cdots+1\cdot x_n$, $[x_k, y_k] := x_1y_1+\cdots+x_ny_n$ 等等.

b) 如果用量 $x^1, \cdots, x^m, y$ 之间的关系式

$$y=\sum_{i=1}^{m} a_i x^i + b$$

代替等式 (26), 试写出这种情况下的最小二乘法原理导出的关于系数 $a_1, \cdots, a_m, b$ 的线性方程组.

c) 如果物理量 $x_1, \cdots, x_n$ 与 y 之间具有下述关系

$$y = cx_1^{\alpha_1}\cdot\cdots\cdot x_n^{\alpha_n},$$

如何利用最小二乘法求它的经验公式?

d) (М.Джермен) 对几十个特殊的环节蠕虫杂色沙蚕在不同温度 T 下测量了它们的心脏的收缩频率 R. 列表如下 (表中频率表示为关于 15°C 时收缩频率的百分数).

温度/°C	频率/%	温度/°C	频率/%
0	39	20	136
5	54	25	182
10	74	30	254
15	100		

R 对 T 的依赖关系与指数关系相像, 可以认为是 $R = Ae^{bT}$, 求常数 A 与 b 的值, 使得上述 R 与 T 的函数关系与实验结果最好地符合.

9. a) 证明: 在例 5 中考察的惠更斯问题中, 变量 $m_1, \cdots, m_n$ 中即使有一个趋于无穷大, 函数 (18) 就趋于零.

b) 证明: 函数 (18) 在 $\mathbb{R}^n$ 中有极大值点, 因此这个函数在 $\mathbb{R}^n$ 中唯一的极值点应当是它的极大值点.

c) 证明: 由公式 (19) 定义的函数 v 是 n 的单调增加函数, 并且求它当 $n\to\infty$ 时的极限.

10. a) 所谓外圆磨具是一个带有粗糙表面的快速旋转的圆形磨具, 它在与慢速旋转 (与它相比较而言) 的圆形零件相接触中扮演着锉的角色 (图 55).

圆 K 逐渐移向零件 Π, 按规定的尺寸吃掉厚为 H 的一层金属并且使得零件表面形成光滑的工作面. 这个曲面在将来的机械中通常容易损坏, 为了延长它的服务期限, 金属零件要通过预先淬火以提高它的强度. 但是因外圆磨与零件接触产生高温, 可能改变了某个 Δ 层金属的结构 (这是经常发生的), 从而降低了这层金属的强度. 量 Δ 单调地依赖于圆 K 移近零件的速度 s, 即 $\Delta=\varphi(s)$. 已经知道, 存在某个临界速度 $s_0>0$, 此时还有 $\Delta=0$, 而当 $s>s_0$ 时, 就有 $\Delta>0$.

为了今后方便, 考察定义在 $\Delta>0$ 范围内的反函数

$$s=\psi(\Delta),$$

这里, ψ 是由实验确定的定义在 $\Delta\geqslant 0$ 范围内的单调增加函数且 $\psi(0)=s_0>0$.

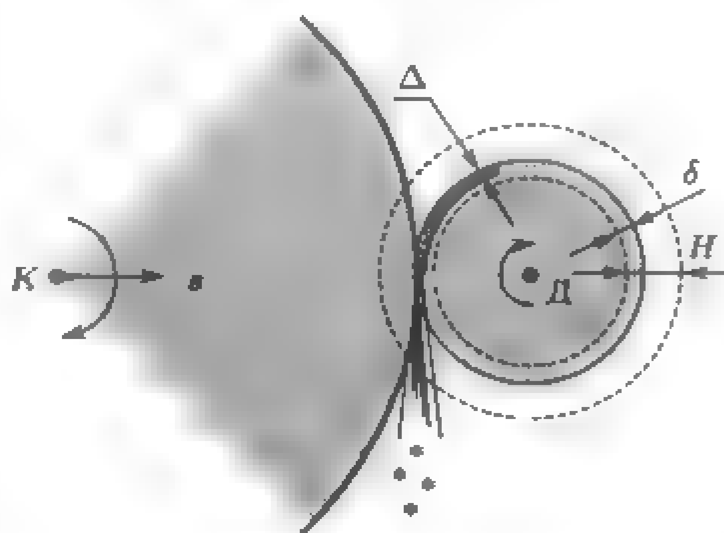


图 55

磨削加的规范应是在最终得到的产品表面金属结构没有变化.

显然, 在上述条件下, 从工作效率考虑, 外圆磨的最优给进速度是

$$s = \psi(\delta)$$

其中 $\delta = \delta(t)$ 是到时刻 t 时还没磨削的金属层厚度, 或者说是在时刻 t, 外圆磨的外缘到最后磨削好的产品表面的距离. 请你解释上面的结论

b) 在最优外圆磨移动速度下, 求磨削厚度为 H 的金属层时所需的时间.

c) 当函数 $\Delta \overset{\psi}{\mapsto} s$ 是线性函数 $s = s_0 + \lambda\Delta$ 时, 求最优外圆磨移动速度 s 与时间 t 的函数关系 $s = s(t)$.

由于某些磨床机构上的特点, 它的速度 s 的变化只能是离散的, 这就产生了一个优化问题: 在允许磨床速度 s 仅仅转换 n 次的条件下, 如何使生产过程获得最好的效率. 对下面问题的回答便可解决这个问题.

d) 在速度 s 按最优方式连续变化的情况下, 你能阐述在 b) 中得到的磨削时间

$$t(H) = \int_0^H \frac{d\delta}{\psi(\delta)}$$

的几何意义吗?

e) 当 s 由最优连续变化转到最优多级变化时, 在时间上的损失应如何作几何解释?

f) 设区间 $[0, H]$ 的点

$$0 = x_{n+1} < x_n < \cdots < x_1 < x_0 = H$$

是速度转换点, 试证它们必须满足条件:

$$\frac{1}{\psi(x_{i+1})} - \frac{1}{\psi(x_i)} = -\left(\frac{1}{\psi}\right)'(x_i)(x_{i+1} - x_i) \quad (i = 1, \cdots, n),$$

从而, 在 $[x_i, x_{i+1}]$ 这一段上, 外圆磨移动的速度为

$$s = \psi(x_{i+1}) \quad (i = 1, \cdots, n)$$

g) 证明: 当 ψ 是线性函数时, 即 $\psi(\Delta) = s_0 + \lambda\Delta$, 问题 f) 中区间 $[0, H]$ 上的点 x_i 的位置应使下列数

$$\frac{s_0}{\lambda} < \frac{s_0}{\lambda} + x_n < \cdots < \frac{s_0}{\lambda} + x_1 < \frac{s_0}{\lambda} + H$$

构成几何级数.

11. a) 证明. 曲线 $\Gamma : I \to \mathbb{R}^m$ 的切线关于 $\mathbb{R}^m$ 中所选取的坐标系是确定不变的.

b) 证明: 函数 $y = f(x^1, \cdots, x^m)$ 图像 S 的切平面关于 $\mathbb{R}^m$ 中所选取的坐标系是确定不变的.

c) 设集合 $S \subset \mathbb{R}^m \times \mathbb{R}^1$ 是函数 $y = f(x^1, \cdots, x^m)$ 在 $\mathbb{R}^m \times \mathbb{R}^1$ 中坐标 $(x^1, \cdots, x^m, y)$ 下的图像, 而且也是函数 $\tilde{y} = \tilde{f}(\tilde{x}_1, \cdots, \tilde{x}^m)$ 在 $\mathbb{R}^m \times \mathbb{R}^1$ 中坐标 $(\tilde{x}^1, \cdots, \tilde{x}^m, \tilde{y})$ 下的图像. 证明:S 的切平面关于 $\mathbb{R}^m \times \mathbb{R}^1$ 中坐标的线性变换是确定不变的.

d) 证明: 拉普拉斯算子$\Delta f = \sum\limits_{i=1}^{m} \dfrac{\partial^2 f}{\partial x^{i2}}(x)$ 关于 $\mathbb{R}^m$ 中坐标的正交变换是确定不变的.

§5. 隐函数定理

1. 问题的提出与启发性想法 在本节中将要证明的隐函数定理无论从本身意义上还是从它众多的结果上来说都是很重要的.

首先说明问题是什么.

例如, 我们假定有在平面 $\mathbb{R}^2$ 上点的坐标 x, y 之间的一个关系式

$$x^2 + y^2 - 1 = 0. \tag{1}$$

平面 $\mathbb{R}^2$ 上满足上述关系式的点的集合是单位圆周 (图 56).

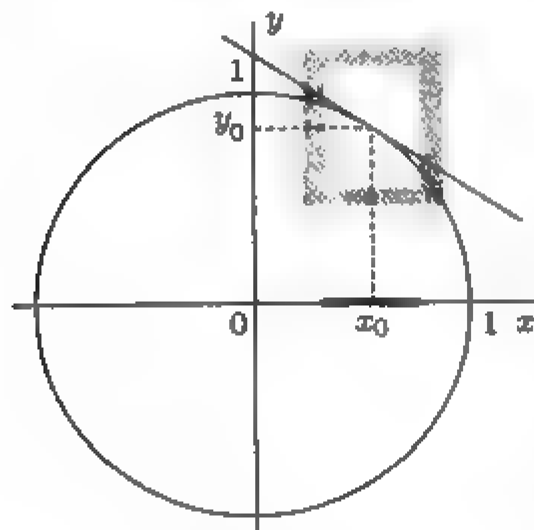


图 56

关系 (1) 告诉我们, 两个坐标中的一个, 例如 x 确定以后, 我们没有权利再任意选取第二个坐标. 由此可见, 关系式 (1) 已经确定了 y 对 x 的依赖性. 我们感兴趣的问题是在什么条件下能从不明显的关系 (1) 中解出明显函数关系 $y = y(x)$.

由方程 (1) 解出 y, 得到

$$y = \pm\sqrt{1 - x^2}, \tag{2}$$

这就是说, 对每一个满足 $|x| < 1$ 的 x, 事实上有两个可容许的 y 值与它对应. 在构造满足关系式 (1) 的函数 $y = y(x)$ 时, 要添加一些补充条件才能从 (2) 式的两个值中确定出一个. 例如, 在 $[-1, 1]$ 的有理点的值是 $\sqrt{1 - x^2}$, 而在无理点的值是 $-\sqrt{1 - x^2}$ 的函数显然满足关系式 (1). 很明显, 改变这个例子中的附加条件, 可以做出满足关系式 (1) 的无穷多个函数.

问题是这样的: 在平面 $\mathbb{R}^2$ 上满足关系式 (1) 的点的集合是某个函数 $y=y(x)$ 的图像吗? 回答显然是否定的, 因为从几何观点看, 它等价于下面的问题: 圆周在某个直线上的正投影与圆周是否存在一一对应?

但是, 我们观察到 (参考图 56) 毕竟在圆周上有单个点 (x_0,y_0) 的邻域内, 它的圆弧与它在 x 轴上的投影之间存在一一对应, 并且可以唯一地表成 $y=y(x)$, 其中 $x\mapsto y(x)$ 是定义在点 x_0 邻域上的连续函数且 $y_0=y(x_0)$. 关于这个问题, 仅仅有两个不好的点 $(-1,0)$ 与 $(1,0)$, 因为把这种点包含在内部的任何一段圆弧都不可能与它们在 x 轴上的投影建立一一对应. 但是, 这些点的邻域在圆周上所处的位置关于 y 轴是好的, 它们是定义在 $y=0$ 的邻域中的连续函数 $x=x(y)$ 的图像, 函数 $x(y)$ 在点 $y=0$ 取值为 1 或 -1, 究竟取 1 还是 -1, 就要看它含有点 $(1,0)$ 还是含有点 $(-1,0)$.

究竟怎样用解析方法得出, 用关系式 (1) 确定的点的几何轨迹, 在它的一点 (x_0,y_0) 的邻域内, 何时能表成显关系式 $y=y(x)$ 或 $x=x(y)$ 呢?

我们将用熟悉的方法进行下面的研究. 考察函数 $F(x,y)=x^2+y^2-1$. 它在点 (x_0,y_0) 邻域内的局部性质可用它的微分

$$F'_x(x_0,y_0)(x-x_0)+F'_y(x_0,y_0)(y-y_0)$$

很好地描述, 这是因为有

$$\begin{aligned}F(x,y)=F(x_0,y_0)+F'_x(x_0,y_0)(x-x_0)+F'_y(x_0,y_0)(y-y_0)\\+o(|x-x_0|+|y-y_0|)\quad (x,y)\to(x_0,y_0).\end{aligned}$$

如果 $F(x_0,y_0)=0$, 并且我们感兴趣的是函数在点 (x_0,y_0) 的邻域内等高集 $F(x,y)=0$ 的性质, 那么, 可根据直线 (切线)

$$F'_x(x_0,y_0)(x-x_0)+F'_y(x_0,y_0)(y-y_0)=0 \tag{3}$$

的位置作出判断.

如果这条直线的位置使得能由它的方程解出 y, 那么, 只要在点 (x_0,y_0) 的邻域内曲线 $F(x,y)=0$ 对这条直线的偏离很小, 就可以期望在点 (x_0,y_0) 的某个邻域内同样可将曲线方程写成 $y=y(x)$ 的形式.

当然, 方程 $F(x,y)=0$ 关于 x 的局部的可解性也是一样的.

对所考察的具体关系式 (1) 写出方程 (3), 得到下面的切线方程

$$x_0(x-x_0)+y_0(y-y_0)=0,$$

当 $y_0\neq 0$ 时, 也就是说, 在圆周 (1) 上除去点 $(-1,0)$ 与 $(1,0)$ 外的所有点 (x_0,y_0) 处, 这个切线方程关于 y 总是可解的. 同样地, 在圆周 (1) 上除去点 $(0,-1)$ 与 $(0,1)$ 外的所有点处, 它关于 x 也是可解的.

2. 隐函数定理的最简单情形 在这一节中, 隐函数定理的证明方法是很直观的, 但不是很有效力的, 因为它仅适合于实变量的实值函数. 在多个关系式的情形, 要用更好的方法, 譬如对其结构作更详细的分析方法得到这个定理, 读者可以参看第十章 (第 II 卷), 本节末的习题 4 也有介绍.

下面的命题是隐函数定理的最简单情形.

命题 1 如果定义在点 $(x_0, y_0) \in \mathbb{R}^2$ 的邻域 $U(x_0, y_0)$ 上的函数 $F: U(x_0, y_0) \to \mathbb{R}$ 满足下列条件

1° $F \in C^{(p)}(U; \mathbb{R})$, 其中 $p \geqslant 1$,

2° $F(x_0, y_0) = 0$,

3° $F'_y(x_0, y_0) \neq 0$,

则存在二维区间 $I = I_x \times I_y \subset U(x_0, y_0)$(其中 $I_x = \{x \in \mathbb{R} || x - x_0| < \alpha\}, I_y = \{y \in \mathbb{R} || y - y_0| < \beta\}$), 函数 $f \in C^{(p)}(I_x, I_y)$, 使对任意的点 $(x, y) \in I_x \times I_y$ 有

$$F(x, y) = 0 \Longleftrightarrow y = f(x), \tag{4}$$

同时, 函数 $y = f(x)$ 在点 $x \in I_x$ 的导数可按下面公式计算:

$$f'(x) = -[F'_y(x, f(x))]^{-1}[F'_x(x, f(x))]. \tag{5}$$

在开始证明之前, 先给出关系式 (4) 的几种可能的表述, 它们一起将使关系式 (4) 本身的意义更加清晰.

命题 1 告诉我们, 在条件 1°,2°,3° 之下, 由关系式 $F(x, y) = 0$ 确定的点的集合在点 (x_0, y_0) 的邻域 $I = I_x \times I_y$ 内是某个 $C^{(p)}(I_x, I_y)$ 类函数 $f: I_x \to I_y$ 的图像.

还可以这样说, 在点 (x_0, y_0) 邻域 I 的范围内可以由方程 $F(x, y) = 0$ 单值地解出 y, 函数 $y = f(x)$ 是这个方程的解, 亦即在 $x \in I_x$ 上, $F(x, f(x)) \equiv 0$. 称 f 是由 $F(x, y) = 0$ 确定的隐函数.

由此也可得到, 如果 $y = \tilde{f}(x)$ 是定义在 I_x 上且满足关系式 $F(x, \tilde{f}(x)) \equiv 0, y_0 = \tilde{f}(x_0)$ 的函数, 并且假定这个函数在点 $x_0 \in I_x$ 连续, 那么, 存在点 x_0 的邻域 $\Delta \subset I_x$, 使得 $\tilde{f}(\Delta) \subset I_y$, 且 $\tilde{f}(x) \equiv f(x), x \in \Delta$.

如果不假定函数 $\tilde{f}$ 在点 x_0 的连续性及 $y_0 = \tilde{f}(x_0)$, 那么, 上述结论不一定正确, 这从上面关于单位圆周的例子中已可看得出来.

现在来证明命题 1.

◀ 为了明确起见, 设 $F'_y(x_0, y_0) > 0$. 因为 $F \in C^{(p)}(U, \mathbb{R})$, 则也在点 (x_0, y_0) 的某个邻域内有 $F'_y(x, y) > 0$. 为了不引进新记号, 不失一般性, 可以认为在原邻域 $U(x_0, y_0)$ 内的一切点 (x, y), 都有 $F'_y(x, y) > 0$.

另外, 如果有必要的话, 先将邻域 $U(x_0, y_0)$ 缩小, 认为它就是中心在点 (x_0, y_0), 半径为某个数 $r = 2\beta > 0$ 的圆.

因为在 $U(x_0,y_0)$ 上有 $F'_y(x,y)>0$, 则关于 y 的函数 $F(x_0,y)$ 在闭区间 $y_0-\beta\leqslant y\leqslant y_0+\beta$ 上定义且单调增加, 因此,

$$F(x_0,y_0-\beta)<F(x_0,y_0)=0<F(x_0,y_0+\beta).$$

由于函数 F 在 U 上连续, 可以找到一个正数 $\alpha<\beta$, 使得当 x 满足不等式 $|x-x_0|\leqslant\alpha$ 时, 有

$$F(x,y_0-\beta)<0<F(x,y_0+\beta).$$

现在证明, 矩形 $I=I_x\times I_y$, 其中

$$I_x=\{x\in\mathbb{R}||x-x_0|<\alpha\}, I_y=\{y\in\mathbb{R}||y-y_0|<\beta\}$$

是所求的二维区间, 对于 I 中一切点, 关系式 (4) 成立.

对于每个 $x\in I_x$, 作端点为 $(x,y_0-\beta)$ 与 $(x,y_0+\beta)$ 的垂直线段, 在这个线段上把 $F(x,y)$ 看成 y 的函数, 它是严格增加的连续函数, 且在线段的端点具有不同符号的数值, 因此, 对于每个 $x\in I_x$, 存在唯一的点 $y(x)\in I_y$, 使得 $F(x,y(x))\equiv 0$, 设 $y(x)=f(x)$, 我们就证明了关系式 (4).

现在证明, $f\in C^{(p)}(I_x;I_y)$.

首先证明函数 f 在点 x_0 连续且 $f(x_0)=y_0$. 显然, 最后的等式可以这样得到: 当 $x=x_0$ 时有唯一的点 $y(x_0)\in I_y$, 使得 $F(x_0,y(x_0))=0$, 再由条件 $F(x_0,y_0)=0$ 得知 $f(x_0)=y_0$.

确定满足 $0<\varepsilon<\beta$ 的数 ε, 我们可以重复函数 $f(x)$ 存在性的证明, 找到满足 $0<\delta<\alpha$ 的 δ, 使得在二维区间 $\tilde{I}=\tilde{I}_x\times\tilde{I}_y$(其中 $\tilde{I}_x=\{x\in\mathbb{R}||x-x_0|<\delta\},\tilde{I}_y=\{y\in\mathbb{R}||y-y_0|<\varepsilon\}$) 上重新找到一个函数 $\tilde{f}:\tilde{I}_x\to\tilde{I}_y$, 满足关系式

$$(F(x,y)=0,(x,y)\in\tilde{I})\Longleftrightarrow(y=\tilde{f}(x),x\in\tilde{I}_x). \tag{6}$$

但是 $\tilde{I}_x\subset I_x,\tilde{I}_y\subset I_y,\tilde{I}\subset I$, 因此, 从 (4) 与 (6) 得到 $\tilde{f}(x)\equiv f(x),x\in\tilde{I}_x\subset I_x$. 这就证明了当 $|x-x_0|<\delta$ 时, 有

$$|f(x)-f(x_0)|=|f(x)-y_0|<\varepsilon.$$

我们已经证明了函数 f 在点 x_0 的连续性, 但是, 曲线 $F(x,y)=0$ 在 I 中任何一点 $(x,y)\in I$ 同样可作为上述论证的出发点, 因为它也满足条件 2°,3°. 这样, 对于在 I 的范围内每个所考察的点 x, 重复上面过程, 由 (4) 重新得到在点 x 邻域内所对应的函数 f 的一部分, 这意味着函数 f 在点 x 连续, 这样一来, 我们证明了 $f\in C(I_x,I_y)$.

现在证明 $f\in C^{(1)}(I_x,I_y)$, 并推导公式 (5).

假定数 Δx 满足 $x+\Delta x\in I_x$, 且 $y=f(x), y+\Delta y=f(x+\Delta x)$. 在区间 I 上对函数 $F(x,y)$ 应用中值定理, 有

$$\begin{aligned}0&=F(x+\Delta x,f(x+\Delta x))-F(x,f(x))\\&=F(x+\Delta x,y+\Delta y)-F(x,y)\\&=F'_x(x+\theta\Delta x,y+\theta\Delta y)\Delta x+F'_y(x+\theta\Delta x,y+\theta\Delta y)\Delta y,\quad 0<\theta<1,\end{aligned}$$

注意到在 I 中有 $F'_y(x,y)\neq 0$, 由此得到

$$\frac{\Delta y}{\Delta x}=-\frac{F'_x(x+\theta\Delta x,y+\theta\Delta y)}{F'_y(x+\theta\Delta x,y+\theta\Delta y)}. \tag{7}$$

因为 $f\in C(I_x,I_y)$, 则当 $\Delta x\to 0$ 时, 这时也有 $\Delta y\to 0$, 再注意到 $F\in C^{(1)}(U,\mathbb{R})$, 令 $\Delta x\to 0$, 在 (7) 式中取极限得到

$$f'(x)=-\frac{F'_x(x,y)}{F'_y(x,y)},$$

其中 $y=f(x)$. 这就证明了公式 (5).

由复合函数的连续性定理, 从公式 (5) 推知 $f\in C^{(1)}(I_x;I_y)$.

如果 $F\in C^{(2)}(U;\mathbb{R})$, 那么, (5) 式右边可以对 x 求导数, 得到

$$f''(x)=-\frac{[F''_{xx}+F''_{xy}\cdot f'(x)]F'_y-F'_x[F''_{xy}+F''_{yy}\cdot f'(x)]}{(F'_y)^2} \tag{5'}$$

其中 $F'_x,F'_y,F''_{xx},F'''_{xy},F''_{yy}$ 是在点 $(x,f(x))$ 的值.

因此, 若 $F\in C^{(2)}(U;\mathbb{R})$, 则 $f\in C^{(2)}(I_x;I_y)$. 因为形如 (5),(5′) 式右边部分中 f 的导数的阶比左边 f 的导数的阶少 1, 因此由归纳法可以得到, 若 $F\in C^{(p)}(U;\mathbb{R})$, 则 $f\in C^{(p)}(I_x;I_y)$. ►

例 1 回过头来考察上边的关系式 (1), 它表示的平面 $\mathbb{R}^2$ 上的圆周, 并用此例验证命题 1.

这时,

$$F(x,y)=x^2+y^2-1,$$

显然 $F\in C^{(\infty)}(\mathbb{R}^2;\mathbb{R})$, 且

$$F'_x(x,y)=2x, F'_y(x,y)=2y,$$

因此, 若 $y\neq 0$, 则 $F'_y(x,y)\neq 0$. 于是, 根据命题 1, 对于圆周上任意一个不同于点 $(-1,0)$ 与 $(1,0)$ 的点 (x_0,y_0), 可以找到一个邻域, 使得圆周落入这个邻域的一段弧可以表成 $y=f(x)$, 直接计算也可确认这一点, 而且 $f(x)=\sqrt{1-x^2}$ 或 $f(x)=-\sqrt{1-x^2}$.

其次, 由命题 1, 有

$$f'(x_0)=-\frac{F'_x(x_0,y_0)}{F'_y(x_0,y_0)}=-\frac{x_0}{y_0}. \tag{8}$$

直接计算, 知

$$\text{若}\quad f(x)=\sqrt{1-x^2},\quad \text{则} f'(x)=-\frac{x}{\sqrt{1-x^2}},$$
$$\text{若}\quad f(x)=-\sqrt{1-x^2},\quad \text{则} f'(x)=\frac{x}{\sqrt{1-x^2}},$$

可以把这两式合并成一个式子

$$f'(x)=-\frac{x}{f(x)}=-\frac{x}{y},$$

由此得到

$$f'(x_0)=-\frac{x_0}{y_0},$$

这与由命题 1 得到的公式 (8) 是一致的.

重要的是, 利用公式 (5) 或 (8) 计算 $f'(x_0)$ 时, 甚至不需要知道显示式 $y=f(x)$, 只要知道 $f(x_0)=y_0$ 就可以了. 而给定条件 $y_0=f(x_0)$, 是选出水平线 $F(x,y)=0$ 中我们企图表成 $y=f(x)$ 的那一段曲线所必须的.

在圆周的例子中可以看出, 仅仅给定坐标 x_0 还不能确定圆弧, 只能再确定了 y_0, 我们才能从两个可能的圆弧中选出一个.

3. 过渡到依赖关系 $F(x^1,\cdots,x^m,y)=0$ 的情形 很容易把命题 1 推广到 $F(x^1,\cdots,x^m,y)=0$ 的情形.

命题 2 *如果定义在点 $(x_0,y_0)=(x_0^1,\cdots,x_0^m,y_0)\in\mathbb{R}^{m+1}$ 的邻域 U 上的函数 $F:U\to\mathbb{R}$ 满足下列条件:*

1° $F\in C^{(p)}(U;\mathbb{R}),p\geqslant 1$,

2° $F(x_0,y_0)=F(x_0^1,\cdots,x_0^m,y_0)=0$,

3° $F'_y(x_0,y_0)=F'_y(x_0^1,\cdots,x_0^m,y_0)\neq 0$,

则存在 $(m+1)$ 维区间 $I=I_x^m\times I_y^1$, 其中

$$I_x^m=\{x=(x^1,\cdots,x^m)\in\mathbb{R}^m\,|\,|x^i-x_0^i|<\alpha^i,i=1,\cdots,m\},$$
$$I_y^1=\{y\in\mathbb{R}\,|\,|y-y_0|<\beta\},$$

有 $I\subset U$, 且存在函数 $f\in C^{(p)}(I_x^m;I_y^1)$, 对任意的点 $(x,y)\in I_x^m\times I_y^1$, 有

$$F(x^1,\cdots,x^m,y)=0\Longleftrightarrow y=f(x^1,\cdots,x^m), \tag{9}$$

同时, 函数 $y=f(x^1,\cdots,x^m)$ 在任一点 $x\in I_x^m$ 的偏导数可按下面公式

$$\frac{\partial f}{\partial x^i}(x)=-[F_y'(x,f(x))]^{-1}[F_{x^i}'(x,f(x))] \tag{10}$$

计算.

◀ 证明区间 $I^{m+1}=I_x^m\times I_y^1$ 与函数 $y=f(x)=f(x^1,\cdots,x^m)$ 的存在性, 以及 f 在 I_x^m 上的连续性可以逐字逐句地重复命题 1 证明中相应的部分, 唯一需要改变的是把 x 改写成 $(x^1,\cdots,x^m)$, 把 α 改写成 $(\alpha^1,\cdots,\alpha^m)$.

现在如果在函数 $F(x^1,\cdots,x^m,y)$ 与 $f(x^1,\cdots,x^m)$ 中除 x^i 与 y 以外的其他变量都暂时固定, 那么, 我们可以把命题 1 的条件中 x 扮演的角色改由 x^i 担任. 由此得到公式 (10) 的正确性. 由这个公式得到 $\dfrac{\partial f}{\partial x^i}\in C(I_x^m;I_y^1)(i=1,\cdots,m)$, 即 $f\in C^{(1)}(I_x^m;I_y^1)$. 如同命题 1 的证明一样, 用归纳法可以证明当 $F\in C^{(p)}(U;\mathbb{R})$ 时, 有 $f\in C^{(p)}(I_x^m;I_y^1)$. ▶

例 2 假定函数 $F:G\to\mathbb{R}$ 定义在区域 $G\subset\mathbb{R}^m$ 上并且属于 $C^{(1)}(G;\mathbb{R})$. 设 $x_0=(x_0^1,\cdots,x_0^m)\in G, F(x_0)=F(x_0^1,\cdots,x_0^m)=0$. 如果 x_0 不是函数 F 的临界点, 那么, 函数 F 在点 x_0 的偏导数中至少有一个不为零, 例如设 $\dfrac{\partial F}{\partial x^m}(x_0)\neq 0$.

这时, 由命题 2 知道, 在点 x_0 的某个邻域中, 满足方程 $F(x^1,\cdots,x^m)=0$ 的 $\mathbb{R}^m$ 中的点集可以表示成在点 $(x_0^1,\cdots,x_0^{m-i})\in\mathbb{R}^{m-1}$ 的一个邻域中定义的某个函数 $x^m=f(x^1,\cdots,x^{m-1}$ 的图像, 这个函数在这个邻域中连续可微且 $f(x_0^1,\cdots,x_0^{m-1})=x_0^m$.

这样, 在函数 F 的非临界点 x_0 的邻域中, 方程

$$F(x^1,\cdots,x^m)=0$$

给出 $(m-1)$ 维曲面.

特别地, 在 $\mathbb{R}^3$ 的情况, 方程

$$F(x,y,z)=0$$

在非临界点 (x_0,y_0,z_0) 的邻域内给出二维曲面, 这个曲面在条件 $\dfrac{\partial F}{\partial z}(x_0,y_0,z_0)\neq 0$ 之下可以局部地表成

$$z=f(x,y).$$

我们知道, 这个函数的图像在点 (x_0,y_0,z_0) 的切平面方程为

$$z-z_0=\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0)+\frac{\partial f}{\partial y}(x_0,y_0)(y-y_0).$$

但是, 按公式 (10) 有

$$\frac{\partial f}{\partial x}(x_0,y_0) = -\frac{F'_x(x_0,y_0,z_0)}{F'_z(x_0,y_0,z_0)},$$
$$\frac{\partial f}{\partial y}(x_0,y_0) = -\frac{F'_y(x_0,y_0,z_0)}{F'_z(x_0,y_0,z_0)},$$

因此, 切平面方程可以改写成

$$F'_x(x_0,y_0,z_0)(x-x_0) + F'_y(x_0,y_0,z_0)(y-y_0) + F'_z(x_0,y_0,z_0)(z-z_0) = 0,$$

它关于变量 x, y, z 是对称的.

类似地, 在一般情况得到 $\mathbb{R}^m$ 中的超平面方程

$$\sum_{i=1}^{m} F'_{x^i}(x_0)(x^i - x_0^i) = 0,$$

它在点 $x_0 = (x_0^1, \cdots, x_0^m)$ 与由方程 $F(x^1, \cdots, x^m) = 0$ 所表示的曲面相切 (当然, 假设有条件 $F(x_0) = 0$ 且 x_0 不是 F 的临界点).

由得到的方程可以看出, 如果在 $\mathbb{R}^m$ 中给定的是欧几里得结构, 那么可以断定, 向量

$$\operatorname{grad} F(x_0) = \left(\frac{\partial F}{\partial x^1}, \cdots, \frac{\partial F}{\partial x^m}\right)(x_0)$$

在点 $x_0 \in \mathbb{R}^m$ 处与函数 F 的 $r-$ 等高集$F(x) = r$ 所表示的曲面正交.

例如, 定义在 $\mathbb{R}^3$ 中的函数

$$F(x,y,z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}$$

的 $r-$ 等高集: 当 $r < 0$ 时是空集; 当 $r = 0$ 时是一点; 当 $r > 0$ 时是椭球

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = r.$$

如果 (x_0, y_0, z_0) 是这个椭球面上的一点, 那么, 根据已证结果, 向量

$$\operatorname{grad} F(x_0,y_0,z_0) = \left(\frac{2x_0}{a^2}, \frac{2y_0}{b^2}, \frac{2z_0}{c^2}\right)$$

在点 (x_0, y_0, z_0) 处垂直于这个椭球面, 且在此点椭球面的切平面方程为

$$\frac{x_0(x-x_0)}{a^2} + \frac{y_0(y-y_0)}{b^2} + \frac{z_0(z-z_0)}{c^2} = 0.$$

由于点 (x_0, y_0, z_0) 在椭球面上, 它又可改写为

$$\frac{x_0 x}{a^2} + \frac{y_0 y}{b^2} + \frac{z_0 z}{c^2} = r.$$

4. 隐函数定理　现在我们转入对一般的方程组

$$\begin{cases} F^1(x^1,\cdots,x^m,y^1,\cdots,y^n)=0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ F^n(x^1,\cdots,x^m,y^1,\cdots,y^n)=0, \end{cases} \tag{11}$$

的讨论, 将关于 $y^1,\cdots,y^n$ 求解它, 也就是说, 寻求局部等价于方程组 (11) 的函数组 (它们是由 (11) 确定的隐函数)

$$\begin{cases} y^1=f^1(x^1,\cdots,x^m), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ y^n=f^n(x^1,\cdots,x^m). \end{cases} \tag{12}$$

为了书写简单、方便和叙述清晰, 记 $x=(x^1,\cdots,x^m), y=(y^1,\cdots,y^n)$, 则 (11) 式左边可记作 $F(x,y)$, (11) 式可写成 $F(x,y)=0$, 而映射 (12) 式可写成 $y=f(x)$.

如果记 $x_0=(x_0^1,\cdots,x_0^m), y_0=(y_0^1,\cdots,y_0^n), \alpha=(\alpha^1,\cdots,\alpha^m), \beta=(\beta^1,\cdots,\beta^n)$, 那么, 可以用记号 $|x-x_0|<\alpha$ 与 $|y-y_0|<\beta$ 分别表示 $|x^i-x_0^i|<\alpha^i(i=1,\cdots,m)$ 与 $|y^j-y_0^j|<\beta^j(j=1,\cdots,n)$.

其次, 记

$$f'(x)=\begin{pmatrix} \dfrac{\partial f^1}{\partial x^1} & \cdots & \dfrac{\partial f^1}{\partial x^m} \\ \vdots & & \vdots \\ \dfrac{\partial f^n}{\partial x^1} & \cdots & \dfrac{\partial f^n}{\partial x^m} \end{pmatrix}(x), \tag{13}$$

$$F_x'(x,y)=\begin{pmatrix} \dfrac{\partial F^1}{\partial x^1} & \cdots & \dfrac{\partial F^1}{\partial x^m} \\ \vdots & & \vdots \\ \dfrac{\partial F^n}{\partial x^1} & \cdots & \dfrac{\partial F^n}{\partial x^m} \end{pmatrix}(x,y), \tag{14}$$

$$F_y'(x,y)=\begin{pmatrix} \dfrac{\partial F^1}{\partial y^1} & \cdots & \dfrac{\partial F^1}{\partial y^n} \\ \vdots & & \vdots \\ \dfrac{\partial F^n}{\partial y^1} & \cdots & \dfrac{\partial F^n}{\partial y^n} \end{pmatrix}(x,y), \tag{15}$$

注意, 矩阵 $F'_y(x,y)$ 是方阵, 从而, 它有逆阵的充要条件是它的行列式不为零. 当 $n=1$ 时, 矩阵 $F'_y(x,y)$ 仅含一个元素, 它的可逆性等价于这个元素不等于零. 矩阵 $F'_y(x,y)$ 的逆阵通常记作 $[F'_y(x,y)]^{-1}$.

现在叙述本节的基本结果.

定理 (隐函数定理) *如果映射 $F: U \to \mathbb{R}^n$ 定义在点 $(x_0,y_0) \in \mathbb{R}^{m+n}$ 的邻域 U 上, 并且满足下述条件*

1° $F \in C^{(p)}(U;\mathbb{R}^n), p \geqslant 1$,

2° $F(x_0,y_0)=0$,

3° $F'_y(x_0,y_0)$ *有逆阵*,

则存在 $(m+n)$ 维区间 $I = I^m_x \times I^n_y \subset U$ (其中

$$I^m_x = \{x \in \mathbb{R}^m \mid |x-x_0| < \alpha\},$$
$$I^n_y = \{y \in \mathbb{R}^n \mid |y-y_0| < \beta\}),$$

并且存在映射 $f \in C^{(p)}(I^m_x; I^n_y)$, 对任意的点 $(x,y) \in I^m_x \times I^n_y$, 有

$$F(x,y) = 0 \Longleftrightarrow y = f(x), \tag{16}$$

同时

$$f'(x) = -[F'_y(x,f(x))]^{-1}[F'_x(x,f(x))]. \tag{17}$$

称 $f: I^m_x \to I^n_y$ 是由 $F(x,y)=0$ 确定的隐函数.

◂ 定理的证明将依靠命题 2 及行列式最简单的性质. 整个证明过程分成几个小段, 采用归纳法论证.

当 $n=1$ 时, 定理归结为命题 2, 因此它是正确的.

假定定理对于 $(n-1)$ 维是正确的, 要证明它对于 n 维也是正确的.

a) 由条件 3° 知矩阵 (15) 的行列式在点 $(x_0,y_0) \in \mathbb{R}^{m+n}$ 的值不为零, 由此知它在点 (x_0,y_0) 的某个邻域内也不为零, 于是, 在这个矩阵最后一行中至少有一个元素不为零, 不妨认为 $\dfrac{\partial F^n}{\partial y^n} \neq 0$, 否则的话, 只要改变一下记号即可.

b) 这时对关系式

$$F^n(x^1,\cdots,x^m,y^1,\cdots,y^n)=0,$$

应用命题 2 可以找到区间 $\tilde{I}^{m+n} = (\tilde{I}^m_x \times \tilde{I}^{n-1}_y) \times I^1_y \subset U$ 和函数 $\tilde{f} \in C^{(p)}(\tilde{I}^m_x \times \tilde{I}^{n-1}_y; I^1_y)$, 使

$$\begin{aligned}
&(F^n(x^1,\cdots,x^m,y^1,\cdots,y^n)=0, (x^1,\cdots,x^m,y^1,\cdots,y^n) \in \tilde{I}^{m+n}) \\
&\Longleftrightarrow (y^n = \tilde{f}(x^1,\cdots,x^m,y^1,\cdots,y^{n-1}), \\
&(x^1,\cdots,x^m) \in \tilde{I}^m_x, (y^1,\cdots,y^{n-1}) \in \tilde{I}^{n-1}_y).
\end{aligned} \tag{18}$$

c) 把得到的 y^n 的表达式 $y^n = \tilde{f}(x^1, \cdots, x^m, y^1, \cdots, y^{n-1})$ 代入方程组 (11) 的前面 $n-1$ 个方程中, 得到 $n-1$ 个关系式

$$\begin{cases} \Phi^1(x^1, \cdots, x^m, y^1, \cdots, y^{n-1}) \\ := F^1(x^1, \cdots, x^m, y^1, \cdots, y^{n-1}, \tilde{f}(x^1, \cdots x^m, y^1, \cdots, y^{n-1})) = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \Phi^{n-1}(x^1, \cdots, x^m, y^1, \cdots, y^{n-1}) \\ := F^{n-1}(x^1, \cdots, x^m, y^1, \cdots, y^{n-1}, \tilde{f}(x^1, \cdots, x^m, y^1, \cdots, y^{n-1})) = 0. \end{cases} \tag{19}$$

因为 $\tilde{f}(x_0^1, \cdots, x_0^m, y_0^1, \cdots, y_0^{n-1}) = y_0^n$ 且 $F^i(x_0, y_0) = 0(i = 1, \cdots, n)$, 所以 $\Phi^i \in C^{(p)}(\tilde{I}_x^m \times \tilde{I}_y^{n-1}; \mathbb{R})$ 且 $\Phi^i(x_0^1, \cdots, x_0^m, y_0^1, \cdots, y_0^{n-1}) = 0(i = 1, \cdots, n-1)$.

由函数 $\Phi^k(k = 1, \cdots, n-1)$ 的定义, 知

$$\frac{\partial \Phi^k}{\partial y^i} = \frac{\partial F^k}{\partial y^i} + \frac{\partial F^k}{\partial y^n} \cdot \frac{\partial \tilde{f}}{\partial y^i} \quad (i, k = 1, \cdots, n-1). \tag{20}$$

还假定

$$\begin{aligned} &\Phi^n(x^1, \cdots, x^m, y^1, \cdots, y^{n-1}) \\ &:= F^n(x^1, \cdots, x^m, y^1, \cdots, y^{n-1}, \tilde{f}(x^1, \cdots, x^m, y^1, \cdots, y^{n-1})), \end{aligned}$$

由 (18) 知在 $\tilde{I}_x^m \times \tilde{I}_y^{n-1}$ 上有 $\Phi^n \equiv 0$, 因此

$$\frac{\partial \Phi^n}{\partial y^i} = \frac{\partial F^n}{\partial y^i} + \frac{\partial F^n}{\partial y^n} \cdot \frac{\partial \tilde{f}}{\partial y^i} \equiv 0 \quad (i = 1, \cdots, n-1). \tag{21}$$

考虑到 (20) 与 (21), 根据行列式的性质可以知道矩阵 (15) 的行列式等于下面矩阵的行列式

$$\begin{pmatrix} \dfrac{\partial F^1}{\partial y^1} + \dfrac{\partial F^1}{\partial y^n} \cdot \dfrac{\partial \tilde{f}}{\partial y^1} & \cdots & \dfrac{\partial F^1}{\partial y^{n-1}} + \dfrac{\partial F^1}{\partial y^n} \cdot \dfrac{\partial \tilde{f}}{\partial y^{n-1}} & \dfrac{\partial F^1}{\partial y^n} \\ \vdots & & \vdots & \\ \dfrac{\partial F^n}{\partial y^1} + \dfrac{\partial F^n}{\partial y^n} \cdot \dfrac{\partial \tilde{f}}{\partial y^1} & \cdots & \dfrac{\partial F^n}{\partial y^{n-1}} + \dfrac{\partial F^n}{\partial y^n} \cdot \dfrac{\partial \tilde{f}}{\partial y^{n-1}} & \dfrac{\partial F^n}{\partial y^n} \end{pmatrix}$$

$$= \begin{pmatrix} \dfrac{\partial \Phi^1}{\partial y^1} & \cdots & \dfrac{\partial \Phi^1}{\partial y^{n-1}} & \dfrac{\partial F^1}{\partial y^n} \\ \vdots & & \vdots & \vdots \\ \dfrac{\partial \Phi^{n-1}}{\partial y^1} & \cdots & \dfrac{\partial \Phi^{n-1}}{\partial y^{n-1}} & \dfrac{\partial F^{n-1}}{\partial y^n} \\ 0 & \cdots & 0 & \dfrac{\partial F^n}{\partial y^n} \end{pmatrix}.$$

由假设, $\dfrac{\partial F^n}{\partial y^n} \neq 0$, 而由条件矩阵 (15) 的行列式不为零, 因此矩阵

$$\begin{pmatrix} \dfrac{\partial \Phi^1}{\partial y^1} & \cdots & \dfrac{\partial \Phi^1}{\partial y^{n-1}} \\ \vdots & & \vdots \\ \dfrac{\partial \Phi^{n-1}}{\partial y^1} & \cdots & \dfrac{\partial \Phi^{n-1}}{\partial y^{n-1}} \end{pmatrix} (x^1, \cdots, x^m, y^1, \cdots, y^{n-1})$$

的行列式在点 $(x_0^1, \cdots, x_0^m, y_0^1, \cdots, y_0^n)$ 的某个邻域内不为零.

按归纳法假定, 存在区间 $I^{m+n-1} = I_x^m \times I_y^{n-1} \subset \tilde{I}_x^m \times \tilde{I}_y^{n-1}$, 它是点 $(x_0^1, \cdots, x_0^m, y_0^1, \cdots, y_0^{n-1}) \in \mathbb{R}^{m+n-1}$ 的邻域, 并且存在映射 $f \in C^{(p)}(I_x^m; I_y^{n-1})$, 使得在区间 $I^{m+n-1} = I_x^m \times I_y^{n-1}$ 上方程组 (19) 等价于

$$\begin{cases} y^1 = f^1(x^1, \cdots, x^m), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & x \in I_x^m. \\ y^{n-1} = f^{n-1}(x^1, \cdots, x^m), \end{cases} \tag{22}$$

d) 因为 $I_y^{n-1} \subset \tilde{I}_y^{n-1}, I_x^m \subset \tilde{I}_x^m$, 那么, 把 (22) 式代入 (18) 式的函数 $y^n = \tilde{f}(x^1, \cdots, x^m, y^1, \cdots, y^{n-1})$ 中, 就得到变量 y^n 对 $(x^1, \cdots, x^m)$ 的依赖关系

$$y^n = f^n(x^1, \cdots, x^m). \tag{23}$$

e) 现在证明, 函数组

$$\begin{cases} y^1 = f^1(x^1, \cdots, x^m), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & x \in I_x^m \\ y^n = f^n(x^1, \cdots, x^m), \end{cases} \tag{24}$$

确定的映射 $f \in C^{(p)}(I_x^m; I_y^n)$(其中 $I_y^n = I_y^{n-1} \times I_y^1$) 在邻域 $I^{m+n} = I_x^m \times I_y^n$ 的范围内等价于方程组 (11).

事实上, 首先我们在 $\tilde{I}^{m+n} = (\tilde{I}_x^m \times \tilde{I}_y^{n-1}) \times I_y^1$ 的范围内把原方程组 (11) 的最后一个方程用与它等价的关系式 $y^n = \tilde{f}(x, y^1, \cdots, y^{n-1})$ 代替 (因有 (18) 式), 得到第二组方程; 在这第二组方程中的前 $n-1$ 个方程中的变量 y^n 用 $\tilde{f}(x, y^1, \cdots, y^{n-1})$ 代替, 得到第三组方程; 这第三组方程中的前 $n-1$ 个方程 (19) , 在 $I_x^m \times I_y^{n-1} \subset \tilde{I}_x^m \times \tilde{I}_y^{n-1}$ 的范围内, 用与之等价的一组关系 (22) 代替. 这样就得到第四组方程. 然后, 我们就能得到在 $I_x^m \times I_y^{n-1} \times I_y^1 = I^{m+n}$ 的范围内与之等价的最终的方程组 (24), 只要在第四组方程的最后一个方程 $y^n = \tilde{f}(x^1, \cdots, x^m, y^1, \cdots, y^{n-1})$ 中, 把变量 $y^1, \cdots, y^{n-1}$ 的表达式 (22) 代入, 从而把它化成 (23) 即可.

f) 为了完成定理的证明, 还需验证公式 (17).

因为在点 (x_0, y_0) 的邻域 $I_x^m \times I_y^n$ 内，方程 (11) 与 (12) 是等价的，所以

$$F(x, f(x)) \equiv 0, \quad x \in I_x^m.$$

写成坐标形式即是，在区域 I_x^m 中成立

$$F^k(x^1, \cdots, x^m, f^1(x^1, \cdots, x^m), \cdots, f^n(x^1, \cdots, x^m)) \equiv 0 \quad (k = 1, \cdots, n). \tag{25}$$

因为 $f \in C^{(p)}(I_x^m; I_y^n), F \in C^{(p)}(U; \mathbb{R}^n)$，这里 $p \geqslant 1$，则 $F(\cdot, f(\cdot)) \in C^{(p)}(I_x^m; \mathbb{R}^n)$，同时，对恒等式 (25) 求偏导数，得到

$$\frac{\partial F^k}{\partial x^i} + \sum_{j=1}^{n} \frac{\partial F^k}{\partial y^j} \cdot \frac{\partial f^j}{\partial x^i} = 0 \quad (k = 1, \cdots, n; i = 1, \cdots, m). \tag{26}$$

显然，关系式 (26) 等价于矩阵等式

$$F'_x(x, y) + F'_y(x, y) \cdot f'(x) = 0,$$

其中 $y = f(x)$.

考虑到矩阵 $F'_y(x, y)$ 在点 (x_0, y_0) 邻域内的可逆性，由上面等式得到

$$f'(x) = -[F'_y(x, f(x))]^{-1}[F'_x(x, f(x))],$$

这就完成了定理的整个证明. ►

练　习

1. 在坐标为 x, y 的平面 $\mathbb{R}^2$上，用关系式 $F(x, y) = 0$ (其中 $F \in C^{(2)}(\mathbb{R}^2; \mathbb{R})$) 给出了一条曲线，设 (x_0, y_0) 位于这条曲线上且是函数 $F(x, y)$ 的非临界点.

   a) 写出这条曲线在点 (x_0, y_0) 的切线方程.

   b) 证明：如果 (x_0, y_0) 是曲线的拐点，则有等式

   $$(F''_{xx}F'^2_y - 2F''_{xy}F'_xF'_y + F''_{yy}F'^2_x)(x_0, y_0) = 0.$$

   c) 求曲线在点 (x_0, y_0) 的曲率公式.

2. ***m* 个变量的函数的勒让德变换**　由变量 $x^1, \cdots, x^m$ 及函数 $f(x^1, \cdots, x^m)$ 变到新变量 $\xi_1, \cdots, \xi_m$ 及函数 $f^*(\xi_1, \cdots, \xi_m)$ 的勒让德变换由下面关系式确定：

$$\begin{cases} \xi_i = \dfrac{\partial f}{\partial x^i}(x^1, \cdots, x^m) \quad (i = 1, \cdots, m), \\ f^*(\xi_1, \cdots, \xi_m) = \displaystyle\sum_{i=1}^{m} \xi_i x^i - f(x^1, \cdots, x^m). \end{cases} \tag{27}$$

a) 给出勒让德变换 (27), 作为从函数 $f(x)$ 图像上的点的坐标 $(x^1,\cdots,x^m,f(x^1,\cdots,x^m))$ 到图像在此点的切平面方程给出的参数 $(\xi_1,\cdots,\xi_m,f^*(\xi_1,\cdots,\xi_m))$, 这样一种转换的几何解释

b) 证明: 若 $f\in C^{(2)}, \det\left(\dfrac{\partial^2 f}{\partial x^i\partial x^j}\right)\neq 0$, 则勒让德变换在局部范围内是存在的.

c) 和一维情况的凸函数定义一样可以定义凸函数 $f(x)=f(x^1,\cdots,x^m)$(现在的 x 理解为向量 $(x^1,\cdots,x^m)\in\mathbb{R}^m$), 证明: 勒让德变换把凸函数变成凸函数.

d) 证明:

$$df^*=\sum_{i=1}^m x^i d\xi_i+\sum_{i=1}^m \xi_i dx^i-df=\sum_{i=1}^m x^i d\xi_i,$$

由此推出勒让德变换是对合的, 即它满足 $(f^*)^*(x)=f(x)$.

e) 利用 d) 把 (27) 式写成关于变量对称的形式

$$\begin{cases} f^*(\xi_1,\cdots,\xi_m)+f(x^1,\cdots,x^m)=\displaystyle\sum_{i=1}^m \xi_i x^i, \\ \xi_i=\dfrac{\partial f}{\partial x^i}(x^1,\cdots,x^m), x^i=\dfrac{\partial f^*}{\partial \xi_i}(\xi_1,\cdots,\xi_m), \end{cases} \tag{28}$$

或者简记作

$$\begin{cases} f^*(\xi)+f(x)=\xi x, \\ \xi=\nabla f(x), x=\nabla f^*(\xi), \end{cases}$$

其中 $\nabla f(x)=\left(\dfrac{\partial f}{\partial x^1},\cdots,\dfrac{\partial f}{\partial x^m}\right)(x), \nabla f^*(\xi)=\left(\dfrac{\partial f^*}{\partial \xi_1},\cdots,\dfrac{\partial f^*}{\partial \xi_m}\right)(\xi),$

$$\xi x=\xi_i x^i=\sum_{i=1}^m \xi_i x^i.$$

f) 函数的二阶偏导数组成的矩阵称为函数在给定点的**黑塞** (Hesse) **矩阵** (有时也把这个矩阵的行列式称为 Hessian).

设 d_{ij} 与 d^*_{ij} 分别表示函数 $f(x)$ 与 $f^*(\xi)$ 的黑塞矩阵

$$\begin{pmatrix} \dfrac{\partial^2 f}{\partial x^1\partial x^1} & \cdots & \dfrac{\partial^2 f}{\partial x^1\partial x^m} \\ \vdots & & \vdots \\ \dfrac{\partial^2 f}{\partial x^m\partial x^1} & \cdots & \dfrac{\partial^2 f}{\partial x^m\partial x^m} \end{pmatrix}(x), \quad \begin{pmatrix} \dfrac{\partial^2 f^*}{\partial \xi_1\partial \xi_1} & \cdots & \dfrac{\partial^2 f^*}{\partial \xi_1\partial \xi_m} \\ \vdots & & \vdots \\ \dfrac{\partial^2 f^*}{\partial \xi_m\partial \xi_1} & \cdots & \dfrac{\partial^2 f^*}{\partial \xi_m\partial \xi_m} \end{pmatrix}(\xi)$$

的元 $\dfrac{\partial^2 f}{\partial x^i\partial x^j}$ 和 $\dfrac{\partial^2 f^*}{\partial \xi_i\partial \xi_j}$ 的代数余子式, d 与 d^* 分别表示矩阵 (d_{ij}) 和 (d^*_{ij}) 的行列式. 如 $d\neq 0$, 证明: $d\cdot d^*=1$, 且

$$\frac{\partial^2 f}{\partial x^i\partial x^j}(x)=\frac{d^*_{ij}}{d^*}(\xi), \quad \frac{\partial^2 f^*}{\partial \xi_i\partial \xi_j}(\xi)=\frac{d_{ij}}{d}(x).$$

g) 肥皂薄膜绷紧在用金属丝做的环形闭路上, 形成所谓*极小曲面*, 它是绷在此闭路上所有曲面中面积最小的曲面.

如果极小曲面用函数 $z=f(x,y)$ 的图像局部地给出, 那么, 函数 f 应当满足下面的极小曲面方程:

$$(1+f_y'^2)f_{xx}''-2f_x'f_y'f_{xy}''+(1+f_x'^2)f_{yy}''=0.$$

证明: 勒让德变换把上述方程变成方程

$$(1+\eta^2)f_{\eta\eta}^{*\prime\prime}+2\xi\eta f_{\xi\eta}^{*\prime\prime}+(1+\xi^2)f_{\xi\xi}^{*\prime\prime}=0.$$

3. 典则变量与哈密顿 (Hamilton) 方程组

a) 在变分法与经典力学基本原理中扮演重要角色的是下面的*欧拉-拉格朗日方程组*

$$\begin{cases}\left(\dfrac{\partial L}{\partial x}+\dfrac{d}{dt}\dfrac{\partial L}{\partial v}\right)(t,x,v)=0,\\ v=\dot{x}(t),\end{cases}\tag{29}$$

其中 $L(t,x,v)$ 是变量 t,x,v 的函数, 其中 t 通常表示时间, x 表示坐标, v 表示速度.

方程组 (29) 由三个变量的两个方程组成, 从方程组 (29) 中通常希望求出关系式 $x=x(t)$ 与 $v=v(t)$, 因为 $v=\dfrac{dx}{dt}$, 所以本质上归结为寻求 $x=x(t)$.

考虑到 $x=x(t),v=v(t)$, 展开导数 $\dfrac{d}{dt}$, 把方程组 (29) 的第一个方程详细地写出来.

b) 证明: 如果用勒让德变换 (参看习题 2)

$$\begin{cases}p=\dfrac{\partial L}{\partial v},\\ H=pv-L,\end{cases}$$

把变量 t,x,v,L 变成*典则变量* t,x,p,H, 也就是用变量 p,H 替换 v,L, 则欧拉–拉格朗日方程组 (29) 就得到了对称形式

$$\dot{p}=-\frac{\partial H}{\partial x},\quad \dot{x}=\frac{\partial H}{\partial p},\tag{30}$$

它称为*哈密顿方程组*.

c) 在多维情况, 即 $L=L(t,x^1,\cdots,x^m,v^1,\cdots,v^m)$, 这时欧拉–拉格朗日方程组有下面形式

$$\begin{cases}\left(\dfrac{\partial L}{\partial x^i}+\dfrac{d}{dt}\dfrac{\partial L}{\partial v^i}\right)(t,x,v)=0,\\ v^i=\dot{x}^i(t),\end{cases}\quad (i=1,\cdots,m)\tag{31}$$

其中 $x=(x^1,\cdots,x^m),v=(v^1,\cdots,v^m)$.

对变量 $v^1,\cdots,v^m,L$ 作勒让德变换, 把变量 $t,x^1,\cdots,x^m,v^1,\cdots,v^m,L$ 变到典则变量 $t,x^1,\cdots,x^m,p_1,\cdots,p_m,H$, 证明: 方程组 (31) 变成下面形式的哈密顿方程组

$$\dot{p}_i=-\frac{\partial H}{\partial x^i},\quad \dot{x}^i=\frac{\partial H}{\partial p_i}\quad (i=1,\cdots,m).\tag{32}$$

4. 隐函数定理 这个问题的解决将给出另一个证明本节基本定理的方法，它可能不怎么直观和有效，但比上边介绍的方法简短.

a) 假设隐函数定理的条件已经满足，且设 $F_y^i(x,y) = \left(\dfrac{\partial F^i}{\partial y^1}, \cdots, \dfrac{\partial F^i}{\partial y^n}\right)(x,y)$ 是矩阵 $F_y'(x,y)$ 的第 i 行.

证明：对于点 (x_0,y_0) 的某个充分小的邻域 $U = I_x^m \times I_y^n$ 内的所有点 (x,y)，由向量 $F_y^i(x,y)$ 组成的矩阵的行列式不为零.

b) 如果对于每个 $x \in I_x^m$，存在点 $y_1, y_2 \in I_y^n$，使得 $F(x,y_1)=0, F(x,y_2)=0$，则对于每个 $i \in \{1,\cdots,n\}$，可以在联结点 (x,y_1) 与点 (x,y_2) 的线段上找到一点 (x,y_i)，使得

$$F_y^i(x,y_i)(y_2-y_1)=0 \quad (i=1,\cdots,n).$$

试证由此可以推出 $y_1=y_2$，也就是说，如果隐函数$f: I_x^m \to I_y^n$ 存在，则它是唯一的.

c) 证明：如果球 $B(y_0;r) \subset I_y^n$，则当 $\| y-y_0 \|_{\mathbb{R}^n} = r > 0$ 时，$F(x_0,y) \neq 0$.

d) 证明：函数 $\| F(x_0,y) \|^2_{\mathbb{R}^n}$ 连续且在球面 $\| y \|_{\mathbb{R}^n} = r$ 上有正的极小值 μ.

e) 证明：存在 $\delta > 0$，使

$\| F(x,y) \|^2_{\mathbb{R}^n} \geqslant \dfrac{1}{2}\mu$，当 $\| x-x_0 \|_{\mathbb{R}^m} < \delta, \| y-y_0 \|_{\mathbb{R}^n} = r$；$\| F(x,y) \|^2_{\mathbb{R}^n} < \dfrac{1}{2}\mu$，当 $\| x-x_0 \|_{\mathbb{R}^m} < \delta, y=y_0$.

f) 任意固定满足 $\| x-x_0 \| < \delta$ 的 x，函数 $\| F(x,y) \|^2_{\mathbb{R}^n}$ 在球 $\| y-y_0 \|_{\mathbb{R}^n} \leqslant r$ 的某个内点 $y=f(x)$ 达到极小值，又因为矩阵 $F_y'(x,f(x))$ 有逆阵，所以 $F(x,f(x))=0$. 这就断定了隐函数$f: B(x_0;\delta) \to B(y_0;r)$ 的存在性.

g) 如果记 $\Delta y = f(x+\Delta x) - f(x)$，则

$$\Delta y = -[\tilde{F}_y']^{-1}[\tilde{F}_x']\Delta x,$$

其中 $\tilde{F}_y'$ 是由行向量 $F_y^i(x_i,y_i)(i=1,\cdots,n)$ 组成的矩阵，而点 (x_i,y_i) 是以点 (x,y) 与点 $(x+\Delta x, y+\Delta y)$ 为端点的线段上的某个点. 对于 $\tilde{F}_x'$ 也有类似的解释.

证明：从上面关系式可推知函数 $y=f(x)$ 的连续性.

h) 证明：

$$f'(x) = -[\tilde{F}_y'(x,f(x))]^{-1}[\tilde{F}_x'(x,f(x))].$$

5. 考察命题：若 $f(x,y,z)=0$，则 $\dfrac{\partial z}{\partial y} \cdot \dfrac{\partial y}{\partial x} \cdot \dfrac{\partial x}{\partial z} = -1$

a) 给这个命题以精确的意义.

b) 对于克拉贝龙定律 $\dfrac{P \cdot V}{T} =$ 常数，验证上述命题的正确性，并且对于一般的三个变量的函数验证它的正确性.

c) 对于 m 个变量之间的关系式 $f(x^1,\cdots,x^m)=0$，写出与上述类似的命题，并验证它的正确性.

6. 试证：方程

$$z^n + c_1 z^{n-1} + \cdots + c_n = 0$$

的根，当它们互不相同时光滑地依赖于方程的系数.

§6. 隐函数定理的一些推论

1. 反函数定理

定义 1 从 $\mathbb{R}^m$ 中的开集 U 到开集 V 上的映射 $f: U \to V$ 称为 $C^{(p)}$ 类微分同胚或 p 级光滑微分同胚 $(p = 0, 1, 2, \cdots)$, 如果

1) $f \in C^{(p)}(U; V)$;

2) f 是双射;

3) $f^{-1} \in C^{(p)}(V; U)$.

$C^{(0)}$ 类微分同胚称为同胚.

通常我们只考察光滑的情况, 即 $p \in \mathbb{N}$ 或 $p = \infty$ 的情况, 并统称为微分同胚.

以下经常用到的定理, 从原则上确立了, 如果一个映射在某点的微分可逆, 则这个映射本身在该点的某一邻域中也可逆.

定理 1 (反函数定理) *如果映射 $f: G \to \mathbb{R}^m$ 定义在区域 $G \subset \mathbb{R}^m$ 上, 且具有下列条件*

1° $f \in C^{(p)}(G; \mathbb{R}^m), p \geqslant 1$,

2° $y_0 = f(x_0), x_0 \in G$,

3° $f'(x_0)$ *有逆阵*,

则存在点 x_0 的邻域 $U(x_0) \subset G$ 和点 y_0 的邻域 $V(y_0)$, 使得 $f: U(x_0) \to V(y_0)$ 是 $C^{(p)}$ 类微分同胚, 且当 $x \in U(x_0), y = f(x) \in V(y_0)$ 时, 有

$$(f^{-1})'(y) = [f'(x)]^{-1},$$

◀ 把关系式

$$y = f(x)$$

改写成

$$F(x, y) = f(x) - y = 0, \tag{1}$$

的形式. 函数 $F(x, y) = f(x) - y$ 定义在点 $(x_0, y_0) \in \mathbb{R}^m \times \mathbb{R}^m$ 的邻域 $G \times \mathbb{R}^m$ 上.

我们希望在点 (x_0, y_0) 的某个邻域内从方程 (1) 中解出 x. 由条件 1°,2°,3° 知映射F 具有以下性质:

$$F \in C^{(p)}(G \times \mathbb{R}^m; \mathbb{R}^m), p \geqslant 1,$$

$$F(x_0, y_0) = 0$$

$$F'_x(x_0, y_0) = f'(x_0)\text{有逆阵}.$$

根据隐函数定理可以找到点 (x_0, y_0) 的邻域 $I_x \times I_y$ 及映射 $g \in C^{(p)}(I_y; I_x)$, 使得对于任意的点 $(x, y) \in I_x \times I_y$, 有

$$f(x) - y = 0 \Longleftrightarrow x = g(y) \tag{2}$$

且

$$g'(y) = -[F'_x(x,y)]^{-1}[F'_y(x,y)].$$

由于 $F'_x(x,y) = f'(x), F'_y(x,y) = -E$, 其中 E 是单位阵, 因此

$$g'(y) = [f'(x)]^{-1}. \tag{3}$$

如果设 $V = I_y, U = g(V)$, 则关系式 (2) 证实了映射$f : U \to V$ 与 $g : V \to U$ 互逆, 即在 V 上有 $g = f^{-1}$.

因为 $V = I_y$, 则 V 是点 y_0 的邻域, 这就是说, 在条件 1°,2°,3° 之下, 区域 G 的内点 x_0 的像 $y_0 = f(x_0)$ 是 G 的像 $f(G)$ 的内点. 由公式 (3) 知矩阵 $g'(y_0)$ 有逆阵, 这就是说, 映射$g : V \to U$ 关于区域 V 与点 $y_0 \in V$ 也具有性质 1°,2° ,3°, 这就证明了 $x_0 = g(y_0)$ 是集合 $U = g(V)$ 的内点.

由公式 (3) 知条件 1°,2°,3° 对任意的点 $y \in V$ 皆满足, 因此任意的点 $x = g(y)$ 都是集合 U 的内点. 于是, U 是点 x_0 在 $\mathbb{R}^m$ 中的开邻域 (显然也是连通的)

到此, 我们验证了映射$f : U \to V$ 满足定义 1 的所有条件和定理 1 的断言. ▶

举一些例子来说明定理 1. 反函数定理常常用于从一个坐标系到另一个坐标系的坐标转换. 最简单情形是在解析几何和线性代数中研究过的坐标变换, 它有下面形式:

$$\begin{pmatrix} y^1 \\ \vdots \\ y^m \end{pmatrix} = \begin{pmatrix} a_1^1 & \cdots & a_m^1 \\ \vdots & & \vdots \\ a_1^m & \cdots & a_m^m \end{pmatrix} \begin{pmatrix} x^1 \\ \vdots \\ x^m \end{pmatrix},$$

或者简记作 $y^j = a_i^j x^i$. 线性变换 $A : \mathbb{R}_x^m \to \mathbb{R}_y^m$ 有定义在整个 $\mathbb{R}_y^m$ 上的逆变换 $A^{-1} : \mathbb{R}_y^m \to \mathbb{R}_x^m$, 当且仅当, 矩阵 (a_i^j) 有逆阵, 即 $\det(a_i^j) \neq 0$.

反函数定理是这个断言的局部化变异形式, 它基于这样一个事实: 光滑映射在一点的小邻域内的性质就像它在这一点的微分那样好.

例 1 极坐标. 从半平面 $\mathbb{R}_+^2 = \{(\rho, \varphi) \in \mathbb{R}^2 | \rho \geqslant 0\}$ 到平面 $\mathbb{R}^2$ 的映射 $f : \mathbb{R}_+^2 \to \mathbb{R}^2$ 由公式

$$\begin{cases} x = \rho \cos\varphi, \\ y = \rho \sin\varphi \end{cases} \tag{4}$$

给出, 见图 57.

不难算出, 这个映射的雅可比行列式等于 ρ, 即在任何点 (ρ, φ)(其中 $\rho > 0$) 的邻域内不为零. 这样, 公式 (4) 局部可逆, 也就是说, 在局部范围内, 数 ρ, φ 可以作为点的新坐标, 代替以前的笛卡儿坐标 x, y.

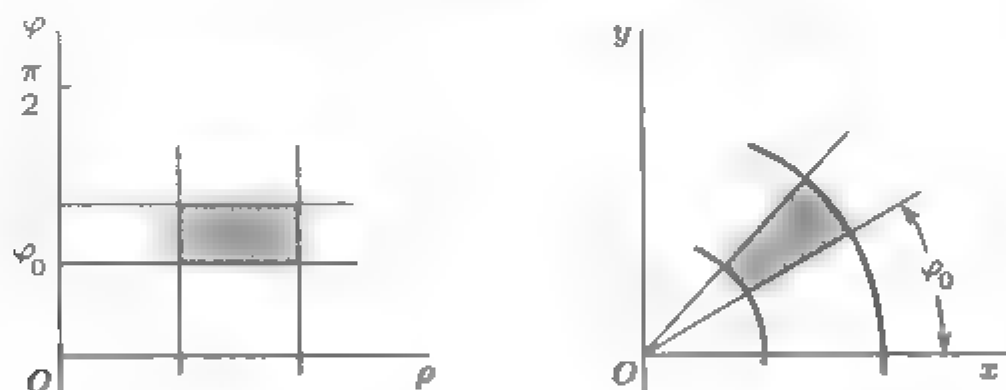


图 57

坐标 (ρ,φ) 是众所周知的平面上的曲线坐标——极坐标，从图 57 可以看出它的几何意义. 注意, 由于函数 $\cos\varphi, \sin\varphi$ 的周期性, 映射 (4) 在 $\rho > 0$ 时仅仅是局部微分同胚的, 而在整个区域上它不是双方单值的, 因此, 从笛卡儿坐标转到极坐标时常常需要选取角度 φ 的分支 (即指出它的变化范围).

在三维空间 $\mathbb{R}^3$ 中的极坐标 (ρ,ψ,φ) 称为球坐标. 它们与笛卡儿坐标之间有下述关系:

$$\begin{cases} z = \rho\ \cos\psi, \\ y = \rho\ \sin\psi\ \sin\varphi, \\ x = \rho\ \sin\psi\ \cos\varphi. \end{cases} \tag{5}$$

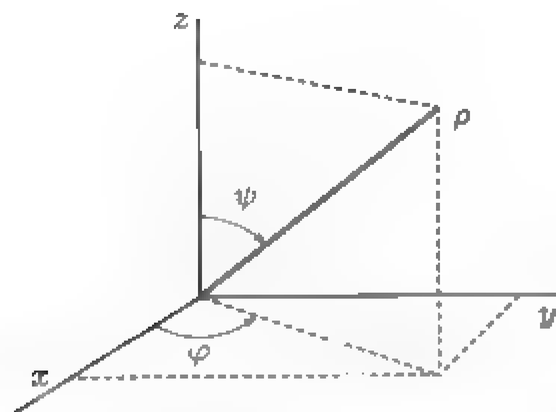


图 58

参数 ρ,ψ,φ 的几何意义如图 58 所示.

这时, 映射 (5) 的雅可比行列式等于 $\rho^2\sin\psi$, 从而, 由定理 1 知, 变换 (5) 在任意一点 (ρ,ψ,φ)(其中 $\rho > 0, \sin\psi \neq 0$) 的邻域内有逆变换.

在空间 (x,y,z) 中, 满足 $\rho =$ 常数, $\varphi =$ 常数, $\psi =$ 常数的点集显然分别是半径为 ρ 的球面, 通过 z 轴的半平面和以 z 轴为对称轴的圆锥面.

这样, 由坐标系 (x,y,z) 转到坐标系 (ρ,ψ,φ) 时, 例如球面和锥面可以局部抻平, 即它们分别对应于小块平面 $\rho =$ 常数与 $\psi =$ 常数. 类似的现象我们在二维情形也观察到了, 平面 (x,y) 上的圆弧对应着平面 (ρ,φ) 上的直线段 (参看图 57). 请注意, 这正是局部的展平.

在 m 维情形, 极坐标由关系式

$$\begin{cases} x^1 = \rho\ \cos\varphi_1, \\ x^2 = \rho\ \sin\varphi_1\cos\varphi_2, \\ \cdots\cdots\ \cdots\cdots\cdots\cdots \\ x^{m-1} = \rho\ \sin\varphi_1\sin\varphi_2\cdot\cdots\cdot\sin\varphi_{m-2}\cos\varphi_{m-1}, \\ x^m = \rho\ \sin\varphi_1\sin\varphi_2\cdot\cdots\cdot\sin\varphi_{m-2}\sin\varphi_{m-1}. \end{cases} \tag{5'}$$

确定. 这个变换的雅可比行列式等于

$$\rho^{m-1}\sin^{m-2}\varphi_1\sin^{m-3}\varphi_2\cdot\cdots\cdot\sin\varphi_{m-2}. \tag{6}$$

由定理 1 知这个变换在它的雅可比行列式不为零的地方也是局部可逆的.

例 2 曲线局部抻直的一般想法. 新坐标的引入通常是为了简化问题所涉及的对象的解析表达式, 使它们在新坐标系中更加易于观察.

例如, 设在平面 $\mathbb{R}^2$ 上用方程

$$F(x,y)=0$$

给出了一条曲线, F 是光滑函数, 点 (x_0,y_0) 在曲线上, 即 $F(x_0,y_0)=0$, 设它不是函数 F 的临界点, 例如设 $F'_y(x_0,y_0)\neq 0$.

我们试图选取坐标 ξ、η, 使得在这个坐标系下, 曲线含点 (x_0,y_0) 的一段弧变成坐标直线, 例如 $\eta=0$ 上的一个线段.

令

$$\begin{cases}\xi=x-x_0,\\ \eta=F(x,y),\end{cases}$$

这个变换的雅可比矩阵

$$\begin{pmatrix}1 & 0\\ F'_x & F'_y\end{pmatrix}(x,y)$$

的行列式是 $F'_y(x,y)$, 据假设条件, 它在点 (x_0,y_0) 不为零, 因此, 由定理 1 知这个映射是从点 (x_0,y_0) 的邻域到点 $(\xi,\eta)=(0,0)$ 的邻域上的微分同胚. 这就是说, 在所指出的邻域范围内, 数 ξ,η 可以作为点 (x_0,y_0) 邻域内的点的新坐标. 在新坐标系下, 我们的曲线方程显然是 $\eta=0$, 在这个意义上我们真地把曲线局部抻直了 (图 59).

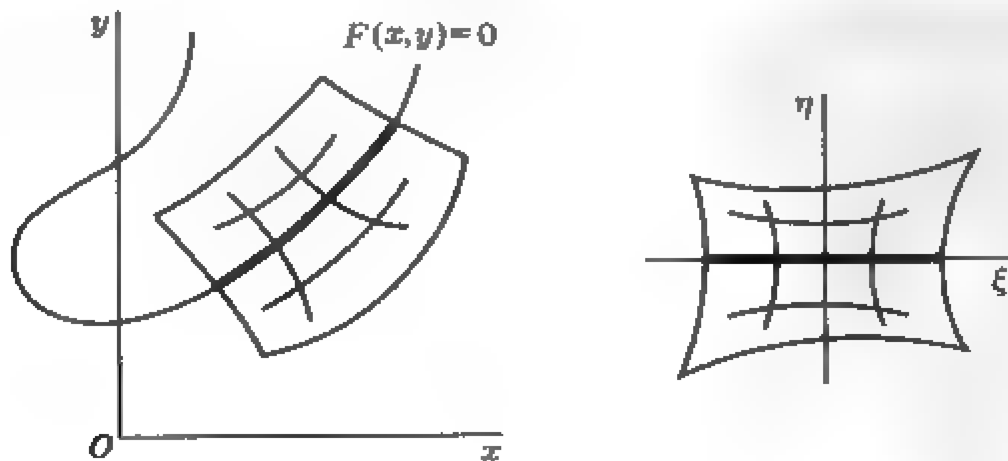


图 59

2. 局部地把光滑映射化为典则形式 我们在这里考察的仅仅是这些类型中的一个问题, 这就是任意具有常数秩的光滑映射, 在适当选取的坐标系下, 均可化为一个确定的典则形式.

我们记得, 区域 $U \subset \mathbb{R}^m$ 上的光滑映射 $f: U \to \mathbb{R}^n$ 在点 $x \in U$ 的秩, 指的是, 它在这点的切映射的秩, 即矩阵 $f'(x)$ 的秩. 映射 f 在点 x 的秩通常记作 $\operatorname{rank} f(x)$.

定理 2 (秩定理) 设 $f: U \to \mathbb{R}^n$ 是定义在点 $x_0 \in \mathbb{R}^m$ 的邻域 $U \subset \mathbb{R}^m$ 上的映射, 如果 $f \in C^{(p)}(U; \mathbb{R}^n), p \geqslant 1$ 且在每点 $x \in U$ 映射 f 有同一秩 k, 则存在点 x_0 与点 $y_0 = f(x_0)$ 的邻域 $O(x_0)$ 与 $O(y_0)$ 以及定义在它们上面的 $C^{(p)}$ 类微分同胚 $u = \varphi(x)$ 与 $v = \psi(y)$, 使得在点 $u_0 = \varphi(x_0)$ 的邻域 $O(u_0) = \varphi(O(x_0))$ 内, 映射 $v = \psi \circ f \circ \varphi^{-1}$ 的坐标表示式如下:

$$(u^1, \cdots, u^k, \cdots, u^m) = u \mapsto v = (v^1, \cdots, v^n) = (u^1, \cdots, u^k, 0, \cdots, 0). \tag{7}$$

换句话说, 定理断定 (图 60), 可以选取坐标 $(u^1, \cdots, u^m)$ 代替 $(x^1, \cdots, x^m)$, 选取坐标 $(v^1, \cdots, v^n)$ 代替坐标 $(y^1, \cdots, y^n)$, 使得我们的映射在这些新坐标下局部地具有 (7), 即 k 秩线性映射的典则形式.

图 60

◀ 把映射 f 写成坐标形式

$$\begin{cases} y^1 = f^1(x^1, \cdots, x^m), \\ \cdots\cdots \\ y^k = f^k(x^1, \cdots, x^m), \\ y^{k+1} = f^{k+1}(x^1, \cdots, x^m), \\ \cdots\cdots \\ y^n = f^n(x^1, \cdots, x^m), \end{cases} \tag{8}$$

为了不改变坐标的编号, 我们认为映射 $f: U \to \mathbb{R}^n_y$ 在任一点 $x \in U$ 的雅可比矩阵的左上角的 k 阶主子式的秩不为零.

考察定义在点 x_0 的邻域 U 上的用下面等式定义的映射:

$$
\begin{cases}
u^1 = \varphi^1(x^1, \cdots, x^m) = f^1(x^1, \cdots, x^m), \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
u^k = \varphi^k(x^1, \cdots, x^m) = f^k(x^1, \cdots, x^m), \\
u^{k+1} = \varphi^{k+1}(x^1, \cdots, x^m) = x^{k+1}, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
u^m = \varphi^m(x^1, \cdots, x^m) = x^m,
\end{cases}
\tag{9}
$$

它的雅可比矩阵是

$$
\left(\begin{array}{ccc|ccc}
\dfrac{\partial f^1}{\partial x^1} & \cdots & \dfrac{\partial f^1}{\partial x^k} & \dfrac{\partial f^1}{\partial x^{k+1}} & \cdots & \dfrac{\partial f^1}{\partial x^m} \\
\vdots & & \vdots & \vdots & & \vdots \\
\dfrac{\partial f^k}{\partial x^1} & \cdots & \dfrac{\partial f^k}{\partial x^k} & \dfrac{\partial f^k}{\partial x^{k+1}} & \cdots & \dfrac{\partial f^k}{\partial x^m} \\
\hline
 & & & 1 & & 0 \\
 & 0 & & & \ddots & \\
 & & & 0 & & 1
\end{array}\right).
$$

由所作的假设知它的行列式在 U 上不为零.

由反函数定理知, 映射 $u = \varphi(x)$ 是把点 x_0 的某个邻域 $\tilde{O}(x_0) \subset U$ 变到点 $u_0 = \varphi(x_0)$ 的邻域 $\tilde{O}(u_0) = \varphi(\tilde{O}(x_0))$ 的 p 级光滑微分同胚.

比较 (8) 与 (9) 式看出, 复合映射 $g = f \circ \varphi^{-1} : \tilde{O}(u_0) \to \mathbb{R}^n_y$ 有下面的坐标表示式

$$
\begin{cases}
y^1 = f^1 \circ \varphi^{-1}(u^1, \cdots, u^m) = u^1, \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
y^k = f^k \circ \varphi^{-1}(u^1, \cdots, u^m) = u^k, \\
y^{k+1} = f^{k+1} \circ \varphi^{-1}(u^1, \cdots, u^m) = g^{k+1}(u^1, \cdots, u^m), \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
y^n = f^n \circ \varphi^{-1}(u^1, \cdots, u^m) = g^n(u^1, \cdots, u^m).
\end{cases}
\tag{10}
$$

因为映射 $\varphi^{-1} : \tilde{O}(u_0) \to \tilde{O}(x_0)$ 在任一点 $u \in \tilde{O}(u_0)$ 有最大秩 m, 而映射 $f : \tilde{O}(x_0) \to \mathbb{R}^n_y$ 在任一点 $x \in \tilde{O}(x_0)$ 的秩为 k , 那么, 由线性代数知道, 矩阵 $g'(u) = f'(\varphi^{-1}(u)) \cdot (\varphi^{-1})'(u)$ 在任一点 $u \in \tilde{O}(u_0)$ 的秩为 k.

直接计算映射 (10) 的雅可比矩阵有

$$\left(\begin{array}{ccc|ccc} 1 & & 0 & & & \\ & \ddots & & & 0 & \\ 0 & & 1 & & & \\ \hline \dfrac{\partial g^{k+1}}{\partial u^1} & \cdots & \dfrac{\partial g^{k+1}}{\partial u^k} & \dfrac{\partial g^{k+1}}{\partial u^{k+1}} & \cdots & \dfrac{\partial g^{k+1}}{\partial u^m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \dfrac{\partial g^n}{\partial u^1} & \cdots & \dfrac{\partial g^n}{\partial u^k} & \dfrac{\partial g^n}{\partial u^{k+1}} & \cdots & \dfrac{\partial g^n}{\partial u^m} \end{array}\right).$$

由此得到, 在任意点 $u \in \tilde{O}(u_0)$ 有 $\dfrac{\partial g^j}{\partial u^i}(u) = 0(i = k+1, \cdots, m; j = k+1, \cdots, n)$. 可以认为邻域 $\tilde{O}(u_0)$ 是凸的 (例如可以缩小 $\tilde{O}(u_0)$ 为中心在 u_0 的球), 由此可以断定, 函数 $g^j(j = k+1, \cdots, n)$ 实际上不依赖于变量 $u^{k+1}, \cdots, u^m$.

于是, 映射 (10) 可以写成

$$\begin{cases} y^1 = u^1, \\ \cdots\cdots \\ y^k = u^k, \\ y^{k+1} = g^{k+1}(u^1, \cdots, u^k), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ y^n = g^n(u^1, \cdots, u^k). \end{cases} \tag{11}$$

现在我们来定义映射 ψ, 令

$$\begin{cases} v^1 = y^1 & =: \psi^1(y), \\ \cdots\cdots & \\ v^k = y^k & =: \psi^k(y), \\ v^{k+1} = y^{k+1} - g^{k+1}(y^1, \cdots, y^k) & =: \psi^{k+1}(y), \\ \cdots\cdots\cdots\cdots\cdots\cdots & \cdots\cdots \\ v^n = y^n - g^n(y^1, \cdots, y^k) & =: \psi^n(y). \end{cases} \tag{12}$$

从函数 $g^j(j = k+1, \cdots, n)$ 的构造可以看出, 映射 ψ 在点 y_0 的某个邻域内有定义, 并且在这个邻域内属于 $C^{(p)}$ 类.

映射 (12) 的雅可比矩阵是

$$\left(\begin{array}{ccc|ccc}1 & & 0 & & & \\ & \ddots & & & 0 & \\ 0 & & 1 & & & \\ \hline -\dfrac{\partial g^{k+1}}{\partial y^1} & \cdots & -\dfrac{\partial g^{k+1}}{\partial y^k} & 1 & & 0 \\ \vdots & & \vdots & & \ddots & \\ -\dfrac{\partial g^n}{\partial y^1} & \cdots & -\dfrac{\partial g^n}{\partial y^k} & 0 & & 1\end{array}\right).$$

它的行列式等于 1, 因此, 由定理 1 知映射 ψ 是从点 $y_0 \in \mathbb{R}^n_y$ 的某个邻域 $\tilde{O}(y_0)$ 到点 $v_0 \in \mathbb{R}^n_v$ 的某个邻域 $\tilde{O}(v_0) = \psi(\tilde{O}(y_0))$ 上的 p 级光滑微分同胚.

比较 (11) 与 (12) 式看出, 当点 u_0 的邻域 $O(u_0) \subset \tilde{O}(u_0)$ 充分小, 以至 $g(O(u_0)) \subset \tilde{O}(y_0)$ 时, 映射 $\psi \circ f \circ \varphi^{-1} : O(u_0) \to \mathbb{R}^n_v$ 是从点 u_0 的邻域 $O(u_0)$ 到点 $v_0 \in \mathbb{R}^n_v$ 的某个邻域 $O(v_0) \subset \tilde{O}(v_0)$ 上的 p 级光滑微分同胚, 并且具有典则形式

$$\begin{cases} v^1 = u^1, \\ \cdots\cdots \\ v^k = u^k, \\ v^{k+1} = 0, \\ \cdots\cdots \\ v^n = 0. \end{cases} \tag{13}$$

令 $\varphi^{-1}(O(u_0)) = O(x_0), \psi^{-1}(O(v_0)) = O(y_0)$, 我们得到定理所要求的点 x_0, y_0 的邻域, 也就完成了定理的证明. ►

定理 2 与定理 1 一样, 显然也是线性代数中相应定理的局部化变异形式.

参考定理 2 的证明, 我们给出下面一些对今后研究有益的附注.

附注 1 如果定理 2 的映射 $f : U \to \mathbb{R}^n$ 在点 x_0 的邻域 $U \subset \mathbb{R}^m$ 中任一点的秩都是 n, 则点 $y_0 = f(x_0)$ 是集合 $f(U)$ 的内点, 即点 y_0 和它的某个邻域都含在 $f(U)$ 中.

◄ 事实上, 映射 $\psi \circ f \circ \varphi^{-1} : O(u_0) \to O(v_0)$ 这时有

$$(u^1, \cdots, u^n, \cdots, u^m) = u \mapsto v = (v^1, \cdots, v^n) = (u^1, \cdots, u^n),$$

的形式, 因此, 点 $u_0 = \varphi(x_0)$ 的邻域的像包含点 $v_0 = \psi \circ f \circ \varphi^{-1}(u_0)$ 的某个邻域.

但是映射 $\varphi : O(x_0) \to O(u_0)$ 与 $\psi : O(y_0) \to O(v_0)$ 都是微分同胚, 因此, 它们把内点变成内点. 现在我们把映射 f 写成 $f = \psi^{-1} \circ (\psi \circ f \circ \varphi^{-1}) \circ \varphi$ 的形式, 可以断定, 点 $y_0 = f(x_0)$ 是点 x_0 的邻域的像的内点. ►

附注 2 如果映射 $f: U \to \mathbb{R}^n$ 在邻域 $U \subset \mathbb{R}^m$ 中任一点的秩 $k < n$, 那么, 根据等式 (8),(12) 和 (13) 在点 $x_0 \in U \subset \mathbb{R}^m$ 的某个邻域内有下面 $n-k$ 个关系式成立

$$f^i(x^1, \cdots, x^m) = g^i(f^1(x^1, \cdots, x^m), \cdots, f^k(x^1, \cdots, x^m))(i = k+1, \cdots, n) \qquad (14)$$

关系式 (14) 显然是关系式 (8),(12),(13) 的推论.

以上所得之关系式是假定矩阵 $f'(x_0)$ 的 k 阶主子式不为零的条件下成立的, 也就是说, 我们要求函数组 $f^1, \cdots, f^k$ 的秩已经是 k. 不然的话, 可以改变函数 $f^1, \cdots, f^n$ 的编号归结为上述情况.

3. 函数相关性

定义 2 连续函数组 $f^i(x^1, \cdots, x^m)(i = 1, \cdots, n)$ 在点 $x_0 = (x_0^1, \cdots, x_0^m)$ 的邻域内称为*函数独立*, 如果对于定义在点

$$y_0 = (y_0^1, \cdots, y_0^n) = (f^1(x_0), \cdots, f^n(x_0)) = f(x_0)$$

邻域上的任何一个连续函数 $F(y^1, \cdots, y^n)$, 仅当 $F(y^1, \cdots, y^n) \equiv 0$ 在点 y_0 邻域内成立时, 关系式

$$F(f^1(x^1, \cdots, x^m), \cdots, f^n(x^1, \cdots, x^m)) \equiv 0$$

才在点 x_0 的邻域内成立.

代数中线性无关的概念是关于线性关系

$$F(y^1, \cdots, y^n) := \lambda_1 y^1 + \cdots + \lambda_n y^n$$

的独立性.

如果函数组不是函数独立的, 则称它*函数相关*.

在代数中, 如果向量组是线性相关的, 那么, 显然其中一个向量必是其他向量的线性组合. 类似地, 对于函数相关的光滑函数组也有相应的结果.

命题 1 *如果函数组 $f^i(x^1, \cdots, x^m)(i = 1, \cdots, n)$ 是在点 $x_0 \in \mathbb{R}^m$ 的邻域 $U(x_0)$ 上定义的一组光滑函数且矩阵*

$$\begin{pmatrix} \dfrac{\partial f^1}{\partial x^1} & \cdots & \dfrac{\partial f^1}{\partial x^m} \\ \vdots & & \vdots \\ \dfrac{\partial f^n}{\partial x^1} & \cdots & \dfrac{\partial f^n}{\partial x^m} \end{pmatrix}(x)$$

在任一点 $x \in U$ 的秩都是 k, 则

a) 当 $k = n$ 时, 函数组在点 x_0 邻域内是函数独立的.

b) 当 $k<n$ 时, 存在点 x_0 邻域及 k 个函数, 设为 $f^1,\cdots,f^k$, 使得其余 $n-k$ 个函数在这个邻域内可以表成如下形式

$$f^i(x^1,\cdots,x^m)=g^i(f^1(x^1,\cdots,x^m),\cdots,f^k(x^1,\cdots,x^m))(i=k+1,\cdots,n)$$

其中 $g^i(y^1,\cdots,y^k)$ 是定义在点 $y_0=(f^1(x_0),\cdots,f^n(x_0))$ 邻域上的光滑函数且仅依赖于动点 $y=(y^1,\cdots,y^n)$ 的 k 个坐标.

◄ 事实上, 如果 $k=n$, 则由秩定理的附注 1 知, 在映射

$$\begin{cases} y^1=f^1(x^1,\cdots,x^m),\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ y^n=f^n(x^1,\cdots,x^m), \end{cases} \tag{15}$$

下, 点 x_0 邻域的像包含点 $y_0=f(x_0)$ 的一个完整的邻域, 因此, 仅当 $F(y^1,\cdots,y^n)\equiv 0$ 在 y_0 邻域内成立时, 关系式

$$F(f^1(x^1,\cdots,x^m),\cdots,f^n(x^1,\cdots,x^m))\equiv 0$$

才在 x_0 邻域内成立, 这就证明了结论 a).

如果 $k<n$ 且假设映射 (15) 的前 k 个函数 $f^1,\cdots,f^k$ 的秩已经是 k, 则由秩定理的附注 2 知存在点 $y_0=f(x_0)$ 的邻域及定义在它上面的 $n-k$ 个函数 $g^i(y)=g^i(y^1,\cdots,y^k)(i=k+1,\cdots,n)$, 它们与给定函数有同级的光滑性, 在点 x_0 某个邻域中关系式 (14) 成立, 这就证明了结论 b). ►

我们证明了, 当 $k<n$ 时, 存在 $n-k$ 个特殊的函数

$$F^i(y)=y^i-g^i(y^1,\cdots,y^k)\quad (i=k+1,\cdots,n),$$

使得在点 x_0 的邻域内函数组 $f^1,\cdots,f^k,\cdots,f^n$ 满足关系式

$$F^i(f^1(x),\cdots,f^k(x),f^i(x))\equiv 0\quad (i=k+1,\cdots,n).$$

4. 局部地分解微分同胚为最简形式的复合 在这里我们要利用反函数定理证明微分同胚可以局部地表示为这样一些微分同胚的复合, 这些微分同胚中的每一个仅仅改变一个坐标.

定义 3 定义在开集 $U\subset\mathbb{R}^m$ 上的微分同胚 $g:U\to\mathbb{R}^m$ 称为最简微分同胚, 如果它的坐标表示式为

$$\begin{cases} y^i=x^i, \quad i\in\{1,\cdots,m\},\ i\neq j,\\ y^j=g^j(x^1,\cdots,x^m), \end{cases}$$

即微分同胚 $g:U\to\mathbb{R}^m$ 仅仅改变被映点坐标中的一个.

命题 2 如果 $f:G\to\mathbb{R}^m$ 是开集 $G\subset\mathbb{R}^m$ 上的微分同胚, 则对任一点 $x_0\in G$, 都存在它的一个邻域, 在其中表示式 $f=g_1\circ\cdots\circ g_m$ 成立, 这里 $g_1,\cdots,g_m$是最简微分同胚.

◀ 采用归纳法, 证明这个命题.

如果映射 f 本身就是最简微分同胚, 那么, 命题是不证自明的.

我们假设命题对于改变的坐标不多于 $k-1$ 个 (其中 $k-1<m$) 的微分同胚是正确的.

现在考察改变 k 个坐标的微分同胚$f:G\to\mathbb{R}^m$:

$$\begin{cases} y^1=f^1(x^1,\cdots,x^m),\\ \cdots\cdots\cdots\cdots\cdots\\ y^k=f^k(x^1,\cdots,x^m),\\ y^{k+1}=x^{k+1},\\ \cdots\cdots\cdots\\ y^m=x^m. \end{cases} \tag{16}$$

我们采取的是改变前 k 个坐标的函数, 但是从将要证明的过程中可以看出, 这仅仅是为了简化写法而不会影响论证的一般性.

因为 f 是微分同胚, 则它的雅可比矩阵 $f'(x)$ 在任一点 $x\in G$ 是非退化的, 这是因为有 $(f^{-1})'(f(x))=[f'(x)]^{-1}$. 固定 $x_0\in G$, 计算矩阵 $f'(x_0)$ 的行列式:

$$\left|\begin{array}{ccc|ccc} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^k} & \frac{\partial f^1}{\partial x^{k+1}} & \cdots & \frac{\partial f^1}{\partial x^m}\\ \vdots & & \vdots & \vdots & & \vdots\\ \frac{\partial f^k}{\partial x^1} & \cdots & \frac{\partial f^k}{\partial x^k} & \frac{\partial f^k}{\partial x^{k+1}} & \cdots & \frac{\partial f^k}{\partial x^m}\\ \hline & & & 1 & & 0\\ & 0 & & & \ddots & \\ & & & 0 & & 1 \end{array}\right|(x_0)=\left|\begin{array}{ccc} \frac{\partial f^1}{\partial x^1} & \cdots & \frac{\partial f^1}{\partial x^k}\\ \vdots & & \vdots\\ \frac{\partial f^k}{\partial x^1} & \cdots & \frac{\partial f^k}{\partial x^k} \end{array}\right|(x_0)\neq 0.$$

这样, 上述的行列式中至少有一个 $k-1$ 阶子式不为零. 为了简化写法, 我们假定它的 $k-1$ 阶主子式不为零. 这时考察用下式定义的辅助映射 $g:G\to\mathbb{R}^m$

$$\begin{cases} u^1=f'(x^1,\cdots,x^m),\\ \cdots\cdots\cdots\cdots\\ u^{k-1}=f^{k-1}(x^1,\cdots,x^m),\\ u^k=x^k,\\ \cdots\cdots\cdots\cdots\\ u^m=x^m, \end{cases} \tag{17}$$

因为映射 $g: G \to \mathbb{R}^m$ 在 $x_0 \in G$ 的雅可比行列式

$$\left|\begin{array}{ccc:ccc} \dfrac{\partial f^1}{\partial x^1} & \cdots & \dfrac{\partial f^1}{\partial x^{k-1}} & \dfrac{\partial f^1}{\partial x^k} & \cdots & \dfrac{\partial f^1}{\partial x^m} \\ \vdots & & \vdots & \vdots & & \vdots \\ \dfrac{\partial f^{k-1}}{\partial x^1} & \cdots & \dfrac{\partial f^{k-1}}{\partial x^{k-1}} & \dfrac{\partial f^{k-1}}{\partial x^k} & \cdots & \dfrac{\partial f^{k-1}}{\partial x^m} \\ \hdashline & & & 1 & & 0 \\ & 0 & & & \ddots & \\ & & & 0 & & 1 \end{array}\right|(x_0) = \left|\begin{array}{ccc} \dfrac{\partial f^1}{\partial x^1} & \cdots & \dfrac{\partial f^1}{\partial x^{k-1}} \\ \vdots & & \vdots \\ \dfrac{\partial f^{k-1}}{\partial x^1} & \cdots & \dfrac{\partial f^{k-1}}{\partial x^{k-1}} \end{array}\right|(x_0) \neq 0,$$

所以, g 在点 x_0 邻域内是微分同胚.

这时在点 $u_0 = g(x_0)$ 的某个邻域内存在 g 的逆映射 $x = g^{-1}(u)$; 于是在点 x_0 的邻域内的任一点 x 都具有新坐标 $(u^1, \cdots, u^m)$.

设 $h = f \circ g^{-1}$. 换句话说, 映射 $y = h(u)$ 就是写成 u 坐标形式的映射 (16)$y = f(x)$. 它作为微分同胚的复合, 在点 u_0 的某个邻域内也是微分同胚. 显然它的坐标表示式为

$$\begin{cases} y^1 = f^1 \circ g^{-1}(u) = u^1, \\ \cdots\cdots\cdots\cdots \\ y^{k-1} = f^{k-1} \circ g^{-1}(u) = u^{k-1}, \\ y^k = f^k \circ g^{-1}(u), \\ y^{k+1} = u^{k+1}, \\ \cdots\cdots\cdots\cdots \\ y^m = u^m, \end{cases}$$

即 h 是最简微分同胚.

但是 $f = h \circ g$, 而由归纳假定知道由公式 (17) 定义的映射 g 可以分解为最简微分同胚 $g_1, \cdots, g_{k-1}$ 的复合, 即 $g = g_1 \circ \cdots \circ g_{k-1}$, 因此, 改变 k 个坐标的微分同胚 f 在点 x_0 的某个邻域内也分解为 k 个最简微分同胚的复合, 即 $f = h \circ g_1 \circ \cdots \circ g_{k-1}$, 由归纳法本命题证毕. ►

5. 莫尔斯[①]引理 列入我们考察范围的还有在非退化临界点的邻域内局部地将光滑实值函数化为典则形式的莫尔斯引理, 这个引理本身非常漂亮, 且在应用中很重要.

定义 4 设 x_0 是定义在点 x_0 的邻域 $U \subset \mathbb{R}^m$ 上属于 $C^{(2)}(U;\mathbb{R})$ 的函数 f 的临界点. 如果 f 在这点的黑塞矩阵的行列式 (即由函数 f 的二阶偏导数组成的矩阵

①莫尔斯 (M. Morse)(1892—1977) —— 美国数学家. 他主要的贡献是把拓扑方法应用于分析学的不同领域.

$\left(\dfrac{\partial^2 f}{\partial x^i \partial x^j}(x_0)\right)$ 的行列式) 不为零, 则称临界点 x_0 是函数 f 的非退化临界点.

如果 x_0 是函数 f 的临界点, 即 $f'(x_0)=0$, 由泰勒公式有

$$f(x)-f(x_0)=\frac{1}{2!}\sum_{i,j}\frac{\partial^2 f}{\partial x^i \partial x^j}(x_0)(x^i-x_0^i)(x^j-x_0^j)+o(\|x-x_0\|^2). \tag{18}$$

莫尔斯引理断定, 在局部范围内可以进行坐标变换 $x=g(y)$, 使得在坐标 y 下函数 f 能写成如下形式

$$f\circ g(y)-f(x_0)=-(y^1)^2-\cdots-(y^k)^2+(y^{k+1})^2+\cdots+(y^m)^2.$$

假如等式 (18) 右边没有余项 $o(\|x-x_0\|^2)$, 那么, 差式 $f(x)-f(x_0)$ 就是通常的二次型, 由代数知道, 线性变换可以把它化为所指出的典则形式. 这样一来, 我们将要证明的命题就是代数中化二次型为典则形式的定理的局部变异形式. 在证明中将会用到证明这个代数定理的思想, 反函数定理以及下面的命题.

阿达马引理 *设 $f:U\to\mathbb{R}$ 是定义在点 $0=(0,\cdots,0)\in\mathbb{R}^m$ 的凸邻域 U 上的 $C^{(p)}(U;\mathbb{R})(p\geqslant 1)$ 类函数且 $f(0)=0$, 则存在函数 $g_i\in C^{(p-1)}(U;\mathbb{R})(i=1,\cdots,m)$, 使得在邻域 U 内下述等式成立*

$$f(x^1,\cdots,x^m)=\sum_{i=1}^m x^i g_i(x^1,\cdots,x^m), \tag{19}$$

且 $g_i(0)=\dfrac{\partial f}{\partial x^i}(0)$.

◀ 等式 (19) 实质上是我们已经知道的具有积分形式余项的泰勒公式 的另外的有用的写法. 它可从等式

$$f(x^1,\cdots,x^m)=\int_0^1\frac{df(tx^1,\cdots,tx^m)}{dt}dt=\sum_{i=1}^m x_i\int_0^1\frac{\partial f}{\partial x^i}(tx^1,\cdots,tx^m)dt,$$

并在其中令

$$g_i(x^1,\cdots,x^m)=\int_0^1\frac{\partial f}{\partial x^i}(tx^1,\cdots,tx^m)dt \quad (i=1,\cdots,m).$$

推出.

显然有 $g_i(0)=\dfrac{\partial f}{\partial x^i}(0)(i=1,\cdots,m)$, 且不难证明 $g_i\in C^{(p-1)}(U;\mathbb{R})$. 但是我们现在不去证明它们, 等将来证明了含参量积分的一般微分法则以后, 从它可以直接推出函数 g_i 具有我们所需要的性质.

这样, 从这个意义上说, 我们证明了阿达马公式 (19). ▶

莫尔斯引理 *如果 $f:G\to\mathbb{R}$ 是定义在开集 $G\subset\mathbb{R}^m$ 上的 $C^{(3)}(G;\mathbb{R})$ 类函数, $x_0\in G$ 是 f 的非退化临界点, 则存在从坐标原点 $0\in\mathbb{R}^m$ 的某个邻域 V 到点*

$x_0 \in \mathbb{R}^m$ 的邻域 U 的微分同胚 $g: V \to U$, 使得对所有的点 $y \in V$, 有

$$(f \circ g)(y) = f(x_0) - [(y^1)^2 + \cdots + (y^k)^2] + [(y^{k+1})^2 + \cdots + (y^m)^2].$$

◀ 用线性变换即可将问题归结为 $x_0 = 0$ 且 $f(x_0) = 0$ 的情形. 下面我们就认为这个条件已经满足.

因为 $x_0 = 0$ 是 f 的临界点, 则公式 (19) 中的 $g_i(0) = 0 (i = 1, \cdots, m)$. 这时, 由阿达马引理有

$$g_i(x^1, \cdots, x^m) = \sum_{j=1}^m x^j h_{ij}(x^1, \cdots, x^m),$$

其中 h_{ij} 是点 0 的邻域上的光滑函数, 于是

$$f(x^1, \cdots, x^m) = \sum_{i,j=1}^m x^i x^j h_{ij}(x^1, \cdots, x^m). \tag{20}$$

如果需要, 在上式中用 $\tilde{h}_{ij} = \frac{1}{2}(h_{ij} + h_{ji})$ 替换 h_{ij}, 因此可以认为 $h_{ij} = h_{ji}$.

我们还注意到, 根据泰勒展开式的唯一性, 由函数 h_{ij} 的连续性得知 $h_{ij}(0) = \frac{\partial^2 f}{\partial x^i \partial x^j}(0)$, 因此, 矩阵 $(h_{ij}(0))$ 是非退化的.

现在我们已经把函数 f 写成类似于二次型的样子, 并且希望把它化为对角线形式.

我们将像在经典情况那样用归纳法证明它.

假设在点 $0 \in \mathbb{R}^m$ 的邻域 U_1 中存在坐标 $u^1, \cdots, u^m$, 即存在用坐标 $u^1, \cdots, u^m$ 表示的微分同胚 $x = \varphi(u)$, 使得

$$f \circ \varphi(u) = \pm(u^1)^2 \pm \cdots \pm (u^{r-1})^2 + \sum_{i,j=r}^m u^i u^j H_{ij}(u^1, \cdots, u^m), \tag{21}$$

其中 $r \geqslant 1$ 且 $H_{ij} = H_{ji}$.

注意, 当 $r = 1$ 时, 令 $H_{ij} = h_{ij}$, 从关系式 (20) 得知 (21) 是成立的.

根据引理的条件, 二次型 $\sum\limits_{i,j=1}^m x^i x^j h_{ij}(0)$ 是非退化的, 即 $\det(h_{ij}(0)) \neq 0$. 变量替换 $x = \varphi(u)$ 是微分同胚, 因此 $\det\varphi'(0) \neq 0$. 但是, 用矩阵 $\varphi'(0)$ 及它的转置阵右乘与左乘二次型 $\sum\limits_{i,j=1}^m x^i x^j h_{ij}(0)$ 的矩阵, 所得到的矩阵是二次型

$$\pm(u^1)^2 \pm \cdots \pm (u^{r-1})^2 + \sum_{i,j=r}^m u^i u^j H_{ij}(0)$$

的矩阵, 它也是非退化的. 因此, 数 $H_{ij}(0) (i, j = r, \cdots, m)$ 之中至少有一个不为零. 可以用线性变换把二次型 $\sum\limits_{i,j=r}^m u^i u^j H_{ij}(0)$ 化为对角线形式, 因此可以认为, 在

等式 (21) 中 $H_{rr}(0) \neq 0$. 由函数 $H_{ij}(u)$ 的连续性知, 在点 $u = 0$ 的某个邻域内有 $H_{rr}(u) \neq 0$.

令 $\psi(u^1, \cdots, u^m) = \sqrt{|H_{rr}(u)|}$. 这时函数 ψ 在点 $u = 0$ 的某个邻域 $U_2 \subset U_1$ 内属于 $C^{(1)}(U_2; \mathbb{R})$ 类. 现在按下式引进新坐标 $(v^1, \cdots, v^m)$:

$$\begin{cases} v^i = u^i, i \neq r, \\ v^r = \psi(u^1, \cdots, u^m)\left[u^r + \sum\limits_{i>r} \dfrac{u^i H_{ir}(u^1, \cdots, u^m)}{H_{rr}(u^1, \cdots, u^m)}\right]. \end{cases} \tag{22}$$

显然, 变换 (22) 在点 $u = 0$ 的雅可比行列式等于 $\psi(0) \neq 0$, 因此, 由反函数定理可以断定, 在点 $u = 0$ 的某个邻域 $U_3 \subset U_2$ 内由 (22) 式定义的映射 $v = \psi(u)$ 是 $C^{(1)}(U_3; \mathbb{R}^m)$ 类微分同胚, 因此, 变量 $(v^1, \cdots, v^m)$ 可以作为 U_3 内的点的坐标.

从等式 (21) 的右边取出所有含 u^r 的项

$$u^r u^r H_{rr}(u^1, \cdots, u^m) + 2 \sum_{j=r+1}^{m} u^r u^j H_{rj}(u^1, \cdots, u^m). \tag{23}$$

在写 (23) 中的和时, 我们用了 $H_{ij} = H_{ji}$.

比较 (22) 与 (23) 可以看出, 表达式 (23) 可以改写成

$$\pm v^r v^r - \frac{1}{H_{rr}(u)}\left[\sum_{i>r} u^i H_{ir}(u_1, \cdots, u_m)\right]^2,$$

在 $v^r v^r$ 前面出现的正负号是由于 $H_{rr} = \pm(\psi)^2$ 的缘故, 并且当 $H_{rr} > 0$ 时取正号, 当 $H_{rr} < 0$ 时取负号.

这样一来, 表达式 (21) 在进行变量替换 $v = \psi(u)$ 以后, 化为等式

$$f \circ \varphi \circ \psi^{-1}(v) = \sum_{i=1}^{r} \pm(v^i)^2 + \sum_{i,j>r} v^i v^j \tilde{H}_{ij}(v^1, \cdots, v^m),$$

其中 $\tilde{H}_{ij}$ 是某个关于指标 i, j 对称的新的光滑函数, 而映射 $\varphi \circ \psi^{-1}$ 是微分同胚. 这样, 我们完成了由 $r - 1$ 到 r 的归纳证明, 从而证明了莫尔斯引理. ►

练　习

1. 计算 $\mathbb{R}^m$ 中由极坐标到笛卡儿坐标的坐标变换 (6) 的雅可比行列式.
2. a) 设 $F: U \to \mathbb{R}$ 是定义在非临界点 $x_0 = (x_0^1, \cdots, x_0^m) \in \mathbb{R}^m$ 邻域 U 上的光滑函数, 证明: 在点 x_0 的某个邻域 $\tilde{U} \subset U$ 中存在曲线坐标 $(\xi^1, \cdots, \xi^m)$, 使得满足条件 $F(x) = F(x_0)$ 的 x 的点集在这个新坐标下用方程 $\xi^m = 0$ 给出.

   b) 设 $\varphi, \psi \in C^{(k)}(D; \mathbb{R})$, 而且, 在区域 D 中成立 $(\varphi(x) = 0) \Rightarrow (\psi(x) = 0)$, 试证, 如果 $\operatorname{grad} \varphi \neq 0$, 则在 D 中成立分解式 $\psi = \theta \cdot \varphi$, 这里 $\theta \in C^{(k-1)}(D; \mathbb{R})$.

3. 设 $f:\mathbb{R}^2\to\mathbb{R}^2$ 是满足柯西 黎曼方程组

$$\frac{\partial f^1}{\partial x^1}=\frac{\partial f^2}{\partial x^2},\frac{\partial f^1}{\partial x^2}=-\frac{\partial f^2}{\partial x^1}$$

的光滑映射.

a) 证明: 这个映射在某点的雅可比行列式等于零的充要条件是 $f'(x)$ 在这点是零矩阵

b) 证明: 如果 $f'(x)\neq 0$, 则 f 在点 x 的邻域内有确定的逆映射 f^{-1} 且满足柯西–黎曼方程组.

4. 函数相关性(直接证明).

a) 证明: 由 $\pi^i(x)=x^i(i=1,\cdots,m)$ 定义的函数组, 在 $\mathbb{R}^m$ 中任一点的邻域内是函数独立的.

b) 证明: 对任何 $f\in C(\mathbb{R}^m;\mathbb{R})$, 函数组 $\pi^1,\cdots,\pi^m,f$ 是函数相关的.

c) 如果由光滑函数组 $f^1,\cdots,f^k(k<m)$ 构造的映射 $f=(f^1,\cdots,f^k)$ 在点 $x_0=(x_0^1,\cdots,x_0^m)\in\mathbb{R}^m$ 的秩是 k, 试证: 在这点的某个邻域内可以将上述函数组扩充为由 m 个光滑函数 $f^1,\cdots,f^m$ 组成的函数独立组.

d) 如果由光滑函数组

$$\xi^i=f^i(x^1,\cdots,x^m)\quad(i=1,\cdots,m)$$

构造的映射 $f=(f^1,\cdots,f^m)$ 在点 $x_0=(x_0^1,\cdots,x_0^m)$ 的秩是 m, 试证: 变量 $(\xi^1,\cdots,\xi^m)$ 在点 x_0 的某个邻域 $U(x_0)$ 内可以作为曲线坐标并且任意函数 $\varphi:U(x_0)\to\mathbb{R}$ 可以写成 $\varphi(x)=F(f^1(x),\cdots,f^m(x))$ 的形式, 其中 $F=\varphi\circ f^{-1}$.

e) 由光滑函数组构造的映射的秩也称为这个函数组的秩. 证明: 如果光滑函数组 $f^i(x^1,\cdots,x^m)(i=1,\cdots,k)$ 的秩等于 k 且函数组 $f^1,\cdots,f^k,\varphi$ 在某点 $x_0\in\mathbb{R}^m$ 的秩也是 k, 则在这点邻域内有 $\varphi(x)=F(f^1(x),\cdots,f^k(x))$. (提示. 利用 c),d) 且证明 $F(f^1,\cdots,f^m)=F(f^1,\cdots,f^k)$.)

5. 证明: 光滑映射 $f:\mathbb{R}^m\to\mathbb{R}^n$ 的秩是下半连续函数, 即在点 $x_0\in\mathbb{R}^m$ 邻域内有 $\operatorname{rank}f(x)\geqslant\operatorname{rank}f(x_0)$.

6. a) 对于函数 $f:\mathbb{R}\to\mathbb{R}$, 直接证明莫尔斯引理.

b) 阐明下面函数在坐标原点是否适合莫尔斯引理:$f(x)=x^3$; $f(x)=x\sin\frac{1}{x}$; $f(x)=e^{-\frac{1}{x^2}}\sin^2\frac{1}{x}$; $f(x,y)=x^3-3xy^2$; $f(x,y)=x^2$.

c) 证明: 函数 $f\in C^{(3)}(\mathbb{R}^m;\mathbb{R})$ 的非退化临界点是孤立的, 也就是说, 对每个这样的点存在一个邻域, 使得除该点本身外, 没有函数 f 其他的临界点.

d) 证明: 在莫尔斯引理中, 函数 f 在非退化临界点邻域中典则形式的负数平方项的个数 k 不依赖于引入变量替换的方式, 也就是说, 不依赖于使函数有典则形式的具体坐标. 这个数 k 称为临界点的指标.

§7. $\mathbb{R}^n$ 中的曲面和条件极值理论

具备 $\mathbb{R}^n$ 空间中关于曲面 (流形) 的某些初步知识, 对于深刻理解在应用中占据重要地位的条件极值理论是很有益的.

1. $\mathbb{R}^n$ 中的 k 维曲面 推广质点运动规律 $x = x(t)$ 的概念, 我们得到 $\mathbb{R}^n$ 中的道路概念, 即定义在区间 $I \subset \mathbb{R}$ 上的连续映射 $\Gamma : I \to \mathbb{R}^n$. 用这个映射的光滑程度来定义道路的光滑程度. 道路的承载子 $\Gamma(I) \subset \mathbb{R}^n$ 可能是 $\mathbb{R}^n$ 中非常复杂的点集, 仅仅有时候可以很勉强地称为曲线, 譬如, 道路的承载子可以是单独的一个点.

类似地, 定义在 k 维区间 $I^k \subset \mathbb{R}^k$ 上的连续或光滑映射 $f : I^k \to \mathbb{R}^n$ 称为 $\mathbb{R}^n$ 中的*k- 道路*, 它的像 $f(I^k)$ 可能完全不是我们想像的 $\mathbb{R}^n$ 中的 k 维曲面, 例如它也可能就是一个点.

设 $f : G \to \mathbb{R}^n$ 是定义在区域 $G \subset \mathbb{R}^k$ 上的光滑映射. 为了使 f 所确定的 $\mathbb{R}^n$ 中的几何图形是用 k 个独立参数 $(t^1, \cdots, t^k) \in G$ 描述的 k 维图形, 由上节知道, 只要 $f : G \to \mathbb{R}^n$ 在每一点 $t \in G$ 的秩等于 k (自然要求 $k \leqslant n$) 即可. 这时, 映射 $f : G \to f(G)$ 在局部范围内 (即在任一点 $t \in G$ 的邻域内) 是相互单值的.

事实上, 设 $\operatorname{rank} f(t_0) = k$, 譬如, 映射 $f : G \to \mathbb{R}^n$ 的坐标表示式

$$\begin{cases} x^1 = f^1(t^1, \cdots, t^k), \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ x^n = f^n(t^1, \cdots, t^k) \end{cases} \tag{1}$$

中前面 k 个函数的秩为 k.

于是由反函数定理知变量 $t^1, \cdots, t^k$ 在点 t_0 的某个邻域 $U(t_0)$ 内可以用变量 $x^1, \cdots, x^k$ 表示, 这就是说, 集合 $f(U(t_0))$ 可以写成

$$x^{k+1} = \varphi^{k+1}(x^1, \cdots, x^k), \cdots, x^n = \varphi^n(x^1, \cdots, x^k)$$

的形式 (亦即, 集合 $f(v(t_0))$ 可相互单值地投影到坐标平面 $x^1, \cdots, x^k$ 上), 因此, 映射 $f : U(t_0) \to f(U(t_0))$ 是相互单值的.

图 61

但是已经有如图 61 所示的一维光滑道路的例子, 明显说明, 从参数区域 G 到空间 $\mathbb{R}^n$ 中的局部相互单值的这样的映射完全不一定在整体上也是相互单值的. 由于轨迹上可能有多重自身相交的点, 因此, 如果我们需要定义 $\mathbb{R}^n$ 中的 k 维光滑曲面是这样一个集合, 它在自己的每一点附近都可由某个 k 维平面 ($\mathbb{R}^n$ 中的 k 维子空间) 块变形而成, 那么, 只把 k 维曲面的典则块 $G \subset \mathbb{R}^k$ 正规地映入 $\mathbb{R}^n$ 中是不够的, 还必须同时注意使它从整体上呈现出是嵌入到这个空间中的.

定义 1 集合 $S \subset \mathbb{R}^n$ 称为空间 $\mathbb{R}^n$ 中的*k 维光滑曲面* (或 $\mathbb{R}^n$ 中的*k 维子流形*), 如果对于每点 $x_0 \in S$, 存在 $\mathbb{R}^n$ 中的邻域 $U(x_0)$ 以及从这个邻域到 $\mathbb{R}^n$ 中的单位正方体

$$I^n = \{t \in \mathbb{R}^n \mid |t^i| \leqslant 1, i = 1, \cdots, n\}$$

上的微分同胚$\varphi: U(x_0) \to I^n$, 使得集合 $S \cap U(x_0)$ 在这个微分同胚下的像在 I^n 中的部分是由关系式 $t^{k+1} = 0, \cdots, t^n = 0$ 确定的 $\mathbb{R}^n$ 中的 k 维平面 (图 62).

图 62

我们将用微分同胚φ 的光滑程度来衡量曲面S 的光滑程度.

如果把变量 $t^1, \cdots, t^n$ 看成邻域 $U(x_0)$ 中点的新坐标, 那么, 定义 1 可以重新简述如下: 集合 $S \subset \mathbb{R}^n$ 称为 $\mathbb{R}^n$ 中的 k 维曲面 (k 维子流形), 如果对于每点 $x_0 \in S$, 可以指出一个邻域 $U(x_0)$ 和这样一组坐标 $t^1, \cdots, t^n$, 使得集合 $S \cap U(x_0)$ 的点能用关系式 $t^{k+1} = \cdots = t^n = 0$ 给出.

定义 1 中单位立方体的角色纯粹是约定的, 它大致上起着标准尺寸或者地图册中一页页地图的作用. 立方体在坐标系 $t^1, \cdots, t^n$ 中的典型位置仅仅涉及区域的标准化问题, 并无更多的意义, 因为 $\mathbb{R}^n$ 中任何一个立方体经过线性同胚总是可以变成标准的单位立方体.

以上这些叙述常常用来简化对一个集合 $S \subset \mathbb{R}^n$ 是 $\mathbb{R}^n$ 中曲面所做的证明.

考察一些例子.

例 1 空间 $\mathbb{R}^n$ 本身是 $C^{(\infty)}$ 类 n 维曲面. 定义中的映射 $\varphi: \mathbb{R}^n \to I^n$ 在这里可取作

$$\xi^i = \frac{2}{\pi} \operatorname{arctg} x^i \quad (i = 1, \cdots, n).$$

例 2 例 1 中的映射同时建立了用条件 $x^{k+1} = \cdots = x^n = 0$ 给出向量空间 $\mathbb{R}^n$ 的子空间, 它是 $\mathbb{R}^n$ 中的 k 维曲面 (或 $\mathbb{R}^n$ 中的 k 维子流形).

例 3 在 $\mathbb{R}^n$ 中给定方程组

$$\begin{cases} a_1^1 x^1 + \cdots + a_k^1 x^k + a_{k+1}^1 x^{k+1} + \cdots + a_n^1 x^n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ a_1^{n-k} x^1 + \cdots + a_k^{n-k} x^k + a_{k+1}^{n-k} x^{k+1} + \cdots + a_n^{n-k} x^n = 0, \end{cases}$$

设它的秩等于 $n-k$, 则由它确定的点集是 $\mathbb{R}^n$ 中的 k 维子流形.

事实上, 例如设行列式

$$\begin{vmatrix} a^1_{k+1} & \cdots & a^1_n \\ \vdots & & \vdots \\ a^{n-k}_{k+1} & \cdots & a^{n-k}_n \end{vmatrix} \neq 0,$$

这时线性变换

$$\begin{cases} t^1 = x^1, \\ \cdots\cdots\cdots\cdots \\ t^k = x^k, \\ t^{k+1} = a^1_1 x^1 + \cdots + a^1_n x^n, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ t^n = a^{n-k}_1 x^1 + \cdots + a^{n-k}_n x^n \end{cases}$$

显然是非退化的. 在坐标系 $t^1, \cdots, t^n$ 中, 我们的集合用条件 $t^{k+1} = \cdots = t^n = 0$ 给出, 这已经在例 2 中考察过了.

例 4 定义在某个区域 $G \subset \mathbb{R}^{n-1}$ 上的光滑函数 $x^n = f(x^1, \cdots, x^{n-1})$ 的图像是 $\mathbb{R}^n$ 中的 $(n-1)$ 维光滑曲面.

事实上, 设

$$\begin{cases} t^i = x^i \quad (i = 1, \cdots, n-1), \\ t^n = x^n - f(x^1, \cdots, x^{n-1}), \end{cases}$$

我们得到坐标系 $t^1, \cdots, t^{n-1}, t^n$, 在这个坐标系中, 函数的图像可用方程 $t^n = 0$ 描绘.

例 5 $\mathbb{R}^2$ 中的圆周 $x^2 + y^2 = 1$ 是 $\mathbb{R}^2$ 中的一维子流形. 这是由上一节得到的直角坐标 (x, y) 到极坐标 (ρ, φ) 的变换的局部可逆性确立的. 在极坐标系中, 上述圆周可用方程 $\rho = 1$ 描绘.

例 6 这个例子是例 3 的推广. 同时从定义 1 看出, 它给出了 $\mathbb{R}^n$ 空间中子流形的一般坐标形式.

设 $F^i(x^1, \cdots, x^n)(i = 1, \cdots, n-k)$ 是光滑函数组, 它的秩是 $n-k$, 我们证明: 由关系式

$$\begin{cases} F^1(x^1, \cdots, x^k, x^{k+1}, \cdots, x^n) = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ F^{n-k}(x^1, \cdots, x^k, x^{k+1}, \cdots, x^n) = 0 \end{cases} \tag{2}$$

确定的集合 S 是 $\mathbb{R}^n$ 中的 k 维子流形.

设在点 $x_0 \in S$ 处有条件

$$\begin{vmatrix} \dfrac{\partial F^1}{\partial x^{k+1}} & \cdots & \dfrac{\partial F^1}{\partial x^n} \\ \vdots & & \vdots \\ \dfrac{\partial F^{n-k}}{\partial x^{k+1}} & \cdots & \dfrac{\partial F^{n-k}}{\partial x^n} \end{vmatrix} (x_0) \neq 0, \tag{3}$$

这时由反函数定理知变换

$$\begin{cases} t^i = x^i & (i = 1, \cdots, k), \\ t^i = F^i(x^1, \cdots, x^n) & (i = k+1, \cdots, n) \end{cases}$$

在点 x_0 的某个邻域内是微分同胚.

在坐标 $t^1, \cdots, t^n$ 中, 原来的方程组有 $t^{k+1} = \cdots = t^n = 0$ 的形式, 因此 S 是 $\mathbb{R}^n$ 中的 k 维光滑曲面.

例 7 平面 $\mathbb{R}^2$ 上满足方程 $x^2 - y^2 = 0$ 的点集 E 是由通过坐标原点的两条直线组成的. 这个集合不是 $\mathbb{R}^2$ 中的一维子流形 (验证!), 因为它有自身相交的点.

如果从 E 中除去坐标原点 $0 \in \mathbb{R}^2$, 那么, 集合 $E \setminus 0$ 显然满足定义 1. 注意, 集合 $E \setminus 0$ 不是连通的, 它由四条没有公共点的射线组成.

这就是说, 满足定义 1 的 $\mathbb{R}^n$ 中的 k 维曲面可以是非连通的子集, 它们由一些连通分支组成 (这些分支是连通的 k 维曲面). $\mathbb{R}^n$ 中的曲面常常理解为连通的 k 维曲面. 对于我们, 在给定的曲面上寻求函数的极值, 将是一个有兴趣的问题. 这是一个局部性问题, 因此, 并不要求曲面连通性这个条件.

例 8 如果定义在区域 $G \subset \mathbb{R}^k$ 上的光滑映射 $f : G \to \mathbb{R}^n$ 的坐标形式为 (1) 式, 并且它在点 $t_0 \in G$ 的秩为 k, 那么, 存在点 t_0 的邻域 $U(t_0) \subset G$, 它的像 $f(U(t_0)) \subset \mathbb{R}^n$ 是 $\mathbb{R}^n$ 中的光滑曲面.

事实上, 如上面已经看到的那样, 在这种情况, 关系式 (1) 在点 $t_0 \in G$ 某个邻域 $U(t_0)$ 内能替换为与它等价的关系式

$$\begin{cases} x^{k+1} = \varphi^{k+1}(x^1, \cdots, x^k), \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x^n = \varphi^n(x^1, \cdots, x^k) \end{cases} \tag{4}$$

(为了书写简单, 我们假定了函数组 $f^1, \cdots, f^k$ 的秩为 k). 设

$$F^i(x^1, \cdots, x^n) = x^{k+i} - \varphi^{k+i}(x^1, \cdots, x^k) (i = 1, \cdots, n-k),$$

于是 (4) 式可写成 (2) 式. 因为 (3) 式满足, 那么, 由例 6 知集合 $f(U(t_0))$ 实际上是 $\mathbb{R}^n$ 中的 k 维光滑曲面.

2. 切空间 当研究质点在 $\mathbb{R}^3$ 中的运动规律 $x = x(t)$ 时, 从关系式

$$x(t) = x(0) + x'(0)t + o(t), \text{当} t \to 0 \text{时} \tag{5}$$

出发, 并且假定 $t = 0$ 不是映射 $\mathbb{R} \ni t \mapsto x(t) \in \mathbb{R}^3$ 的临界点, 即 $x'(0) \neq 0$, 我们定义了运动轨迹在点 $x(0)$ 的切线, 它是 $\mathbb{R}^3$ 中由参数方程

$$x - x_0 = x'(0)t \tag{6}$$

或

$$x - x_0 = \xi \cdot t \tag{7}$$

确定的线性子集, 其中 $x_0 = x(0)$ 而 $\xi = x'(0)$ 是直线的方向向量.

实际上, 类似的事情我们在 $\mathbb{R}^3$ 中定义函数 $z = f(x, y)$ 图像的切平面时已经做过. 事实上, 在明显的等式 $x = x, y = y$ 中补充一个关系式 $z = f(x, y)$, 我们得到映射

$$\mathbb{R}^2 \ni (x, y) \mapsto (x, y, f(x, y)) \in \mathbb{R}^3,$$

这个映射在点 (x_0, y_0) 的切映射是线性映射

$$\begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ f'_x(x_0, y_0) & f'_y(x_0, y_0) \end{pmatrix} \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}, \tag{8}$$

其中 $z_0 = f(x_0, y_0)$.

如果把 $\begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix}$ 看成 $x - x_0$, 把 $\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$ 看成 t, 用 $x'(0)$ 表示 (8) 式中所考察的映射的雅可比矩阵, 我们注意到它的秩等于 2, 而且, 在这些记号下关系式 (8) 就具有 (6) 的形式了.

关系式 (8) 的特点是它由三个等式

$$\begin{cases} x - x_0 = x - x_0, \\ y - y_0 = y - y_0, \\ z - z_0 = f'_x(x_0, y_0)(x - x_0) + f'_y(x_0, y_0)(y - y_0) \end{cases} \tag{9}$$

组成, 它的总体等价于其最后一个非平凡的等式. 因此, 正是它作为函数 $z = f(x, y)$ 图像上过点 (x_0, y_0, z_0) 的切平面方程留给了我们.

现在可以利用以上所做的观察来定义 $\mathbb{R}^n$ 中 k 维光滑曲面 $S \subset \mathbb{R}^n$ 的 k 维切平面.

从曲面的定义 1 看出, 在 k 维曲面 S 上任一点 $x_0 \in S$ 邻域内的点可以用参数给出, 即利用映射

$$I^k \ni (t^1, \cdots, t^k) \mapsto (x^1, \cdots, x^n) \in S.$$

作为这样的映射可以取 $\varphi^{-1} : I^n \to U(x_0)$ 在 k 维平面 $t^{k+1} = \cdots = t^n = 0$ 上的限制 (参考图 62).

因为 φ^{-1} 是微分同胚, 那么, 映射$\varphi^{-1} : I^n \to U(x_0)$ 在立方体 I^n 中任一点的雅可比行列式不为零. 但是, 这时, 由 φ^{-1} 在所指出的平面上的限制得到的映射

$$I^k \ni (t^1, \cdots, t^k) \mapsto (x^1, \cdots, x^n) \in S$$

在立方体 I^k 中任一点的秩应当是 k.

现在设 $t = (t^1, \cdots, t^k) \in I^k$ 且用 $x = x(t)$ 表示映射 $I^k \ni t \mapsto x \in S$, 我们得到曲面 S 的局部参数表示式, 它具有由等式 (5) 表示的性质, 根据它, 我们取 (6) 式作为切空间的方程或曲面$S \subset \mathbb{R}^n$ 在点 $x_0 \in S$ 的切平面方程.

于是我们采用下面的

定义 2 如果 k 维曲面 $S \subset \mathbb{R}^n (1 \leqslant k \leqslant n)$ 在点 $x_0 \in S$ 的邻域内借助参数形式的光滑映射 $(t^1, \cdots, t^k) = t \mapsto x = (x^1, \cdots, x^n)$ 给出, $x_0 = x(0)$ 且矩阵 $x'(0)$ 的秩为 k, 那么, 称在 $\mathbb{R}^n$ 中用矩阵形式的参数式 $x - x_0 = x'(0)t$ 给出的 k 维平面为*曲面 S 在点 $x_0 \in S$ 的切平面或切空间*.

矩阵等式 (6) 可写成坐标形式

$$\begin{cases} x^1 - x_0^1 = \dfrac{\partial x^1}{\partial t^1}(0)t^1 + \cdots + \dfrac{\partial x^1}{\partial t^k}(0)t^k, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x^n - x_0^n = \dfrac{\partial x^n}{\partial t^1}(0)t^1 + \cdots + \dfrac{\partial x^n}{\partial t^k}(0)t^k. \end{cases} \tag{10}$$

曲面 S 在 $x \in S$ 的切空间像以前一样用记号 TS_x 表示[①].

读者可以独立地完成一个重要而有益的练习, 这就是证明切空间定义的不变性, 并验证, 线性映射 $t \mapsto x'(0)t$, 它是局部定义曲面S 的映射 $t \mapsto x(t)$ 的切映射, 把空间 $\mathbb{R}^k = T\mathbb{R}^k_0$ 映成平面 $TS_{x(0)}$(参看本节末的习题 3).

现在我们来阐明 $\mathbb{R}^n$ 中用方程组 (2) 给出的 k 维曲面S 的切平面方程是什么样子. 为了确定起见, 我们假定在考察的点 $x_0 \in S$ 的邻域内条件 (3) 是满足的.

令

$$(x^1, \cdots, x^k) = u, (x^{k+1}, \cdots, x^n) = v, (F^1, \cdots, F^{n-k}) = F,$$

则方程组 (2) 可写成

$$F(u, v) = 0, \tag{11}$$

①这个记号与最通用的记号 T_xS 或 $T_x(S)$ 的差别是不大的.

而条件 (3) 可写成

$$\det F'_v(u,v) \neq 0. \tag{12}$$

在点 $(u_0,v_0) = (x_0^1,\cdots,x_0^k,x_0^{k+1},\cdots,x_0^n)$ 的邻域内, 由隐函数定理知可把关系式 (11) 化成与它等价的关系式

$$v = f(u). \tag{13}$$

对 (13) 补充恒等式 $u=u$, 我们得到曲面 S 在点 $x_0 \in S$ 邻域内的参数表示式

$$\begin{cases} u = u, \\ v = f(u). \end{cases} \tag{14}$$

根据定义 2, 从 (14) 式得到切平面的参数方程

$$\begin{cases} u - u_0 = E \cdot t, \\ v - v_0 = f'(u_0) \cdot t, \end{cases} \tag{15}$$

其中 E 是单位阵, 而 $t = u - u_0$.

和方程组 (9) 的情形类似, 我们仅留下方程组 (15) 中的非平凡方程

$$v - v_0 = f'(u_0)(u - u_0), \tag{16}$$

它刻画出了切空间变量 $x^1,\cdots,x^k$ 与变量 $x^{k+1},\cdots,x^n$ 之间的联系.

利用隐函数定理有

$$f'(u_0) = -[F'_v(u_0,v_0)]^{-1}[F'_u(u_0,v_0)],$$

于是 (16) 式可写成

$$F'_u(u_0,v_0)(u-u_0) + F'_v(u_0,v_0)(v-v_0) = 0.$$

由此, 重新回到变量 $(x^1,\cdots,x^n) = x$, 就得到要求的切空间$TS_{x_0} \subset \mathbb{R}^n$ 的方程

$$F'_x(x_0)(x - x_0) = 0. \tag{17}$$

与方程 (17) 等价的坐标形式为

$$\begin{cases} \dfrac{\partial F^1}{\partial x^1}(x_0)(x^1 - x_0^1) + \cdots + \dfrac{\partial F^1}{\partial x^n}(x_0)(x^n - x_0^n) = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{\partial F^{n-k}}{\partial x^1}(x_0)(x^1 - x_0^1) + \cdots + \dfrac{\partial F^{n-k}}{\partial x^n}(x_0)(x^n - x_0^n) = 0. \end{cases} \tag{18}$$

由条件知这个方程组的秩是 $n-k$, 因此它给出 $\mathbb{R}^n$ 中的 k 维平面.

仿射方程 (17) 等价于 (如果指定了点 x_0) 向量方程

$$F'_x(x_0)\cdot\xi=0, \tag{19}$$

其中 $\xi=x-x_0$.

这就是说, 如果曲面S 由方程 $F(x)=0$ 定义, 而 $x_0\in S$, 则向量 ξ 在曲面 S 的切平面TS_{x_0} 上, 当且仅当, 它满足条件 (19). 所以, 可以把 TS_{x_0} 看成是满足条件 (19) 的向量 ξ 构成的向量空间.

正是这个缘故, 我们采用了切空间这个术语.

现在来证明下面的命题, 我们已经遇到过它的特殊情形 (参看 §4 第 6 段).

命题 *光滑曲面 $S\subset\mathbb{R}^n$ 在点 $x_0\in S$ 的切空间TS_{x_0} 是由曲面 S 上过点 x_0 的光滑曲线在点 x_0 的切向量组成的.*

◀ 设曲面 S 在点 $x_0\in S$ 的邻域内用方程组 (2) 表示, 它可用简短形式写成

$$F(x)=0, \tag{20}$$

其中 $F=(F^1,\cdots,F^{n-k}), x=(x^1,\cdots,x^n)$. 设 $\Gamma:I\to S$ 是任何一条承载于在曲面 S 上的光滑道路. 令 $I=\{t\in\mathbb{R}||t|<1\}, x(0)=x_0$. 因为 $t\in I$ 时, $x(t)\in S$, 所以在方程 (20) 中代入 $x(t)$ 以后, 得到

$$F(x(t))\equiv0,\quad t\in I, \tag{21}$$

关于 t 微分这个恒等式, 得到

$$F'_x(x(t))\cdot x'(t)\equiv0.$$

特别地, 当 $t=0$ 时, 令 $\xi=x'(0)$, 则有

$$F'_x(x_0)\cdot\xi=0,$$

这就是说, 在点 x_0(时刻 $t=0$) 与轨道相切的向量 ξ 满足切空间TS_{x_0} 的方程 (19).

现在证明, 对任何一个满足方程 (19) 的向量 ξ, 存在一条光滑道路 $\Gamma:I\to S$, 它给出一条曲面S 上的曲线, 它在 $t=0$ 时通过点 x_0, 而且向量 ξ 是它在时刻 $t=0$ 的速度向量.

综合在一起就确立了通过曲面上点 x_0 的光滑曲线的存在性. 在已经作过的命题证明的第一部分中, 我们已经隐含地这样假定.

为确定起见, 我们假设条件 (3) 是满足的. 这时, 我们从 (19) 式 (或与之等价的 (18) 式) 知道由向量 $\xi=(\xi^1,\cdots,\xi^k,\xi^{k+1},\cdots,\xi^n)$ 的前 k 个坐标 $\xi^1,\cdots,\xi^k$ 唯一地确定其余的坐标 $\xi^{k+1},\cdots,\xi^n$, 因此, 如果对于向量 $\tilde\xi=(\xi^1,\cdots,\xi^k,\tilde\xi^{k+1},\cdots,\tilde\xi^n)$ 能证明它满足方程组 (19), 那么, 可以断定 $\tilde\xi=\xi$. 我们将利用这一点.

如上, 为了方便我们仍引用记号 $u=(x^1,\cdots,x^k), v=(x^{k+1},\cdots,x^n), x=(x^1,\cdots,x^n)=(u,v), F(x)=F(u,v)$. 这时方程 (20) 写成 (11), 条件 (3) 写成 (12), 在变量 $x^1,\cdots,x^k$ 的子空间 $\mathbb{R}^k\subset\mathbb{R}^n$ 中, 取直线的参数式

$$\begin{cases} x^1-x_0^1=\xi^1 t, \\ \cdots\cdots\cdots\cdots\cdots \\ x^k-x_0^k=\xi^k t, \end{cases} \quad t\in\mathbb{R},$$

其方向向量为 $(\xi^1,\cdots,\xi^k)$, 用 ξ_u 表示. 这条直线方程可简记成

$$u=u_0+\xi_u t. \tag{22}$$

由隐函数定理可从 (11) 解出 v, 得到光滑函数 (13). 把其中自变量用等式 (22) 的右边部分代入, 考虑到 (22) 式本身, 得到用下式给出的 $\mathbb{R}^n$ 中的光滑曲线

$$\begin{cases} u=u_0+\xi_u t, \\ v=f(u_0+\xi_u t), \end{cases} \quad t\in U(0)\subset\mathbb{R}. \tag{23}$$

因为 $F(u,f(u))\equiv 0$, 显然这条曲线在曲面S 上. 此外, 从等式 (23) 看出, 当 $t=0$ 时, 曲线通过点 $(u_0,v_0)=(x_0^1,\cdots,x_0^k,x_0^{k+1},\cdots,x_0^n)=x_0\in S$.

在 $t=0$, 关于 t 微分恒等式

$$F(u(t),v(t))=F(u_0+\xi_u t, f(u_0+\xi_u t))\equiv 0$$

得到

$$F'_u(u_0,v_0)\xi_u+F'_v(u_0,v_0)\tilde{\xi}_v=0,$$

其中 $\tilde{\xi}_v=v'(0)=(\tilde{\xi}^{k+1},\cdots,\tilde{\xi}^n)$. 这个等式表明, 向量 $\tilde{\xi}=(\xi_u,\tilde{\xi}_v)=(\xi^1,\cdots,\xi^k,\tilde{\xi}^{k+1},\cdots,\tilde{\xi}^n)$ 满足条件 (19), 于是, 由上面已做的说明知道 $\xi=\tilde{\xi}$, 但向量 $\tilde{\xi}$ 是轨线 (23) 在 $t=0$ 时的速度向量, 这就完全证明了本命题. ►

3. 条件极值

a. 问题的提出 最卓越和最受欢迎的微分学成果之一是它所提供的寻求函数极值的方法. 我们已经利用泰勒公式得到的极值的必要条件与充分性微分判别法, 如已见到的那样, 都是关于内部极值点的.

换句话说, 这些结果仅仅适用于研究函数 $\mathbb{R}^n\ni x\mapsto f(x)\in\mathbb{R}$ 在点 $x_0\in\mathbb{R}^n$ 的邻域内的性质, 这时自变量 x 可以取邻域中任何值.

常常出现更加复杂并且从应用的观点看甚至是更加有兴趣的情形, 这就是寻求当自变量变化范围被约束在某些条件下函数的极值. 等周问题就是一个典型的例子, 即求物体在其表面面积固定的条件下, 所具有的最大体积. 为了得到使我们易于了

解的这类问题的数学形式, 我们简化这个问题的提法, 设问题是: 在具有给定周长 $2p$ 的所有长方形中, 求面积 σ 为最大的长方形. 用 x, y 记长方形的边长, 则

$$\sigma(x,y) = xy,$$
$$x + y = p.$$

因此, 应当寻求函数 $\sigma(x,y)$ 在自变量 x, y 满足关系式 $x+y=p$ 的条件下的极值. 也就是说, 我们所寻求的函数极值仅仅是函数在平面 $\mathbb{R}^2$ 上满足所指出关系式的点集上的极值. 当然, 这个具体问题是不难解决的, 只要解出 $y = p - x$, 再代入函数 $\sigma(x,y)$ 的表达式之中, 然后用通常的方法求函数 $x(p-x)$ 的极大值即可. 这个例子仅仅是为了阐明问题的提法而引用的.

在一般情形, 条件极值问题通常是这样的: 求 n 个变量的实值函数

$$y = f(x^1, \cdots, x^n) \tag{24}$$

在这些变量满足方程组

$$\begin{cases} F^1(x^1, \cdots, x^n) = 0 \\ \cdots\cdots\cdots\cdots\cdots\cdots \\ F^m(x^1, \cdots, x^n) = 0 \end{cases} \tag{25}$$

的条件下的极值.

因为我们打算获得极值的微分条件, 所以应假设所考察的所有函数都是可微的甚至是连续可微的. 如果函数组 $F^1, \cdots, F^m$ 的秩是 k, 那么, 条件 (25) 给出 $\mathbb{R}^n$ 中某个 k 维光滑曲面S, 从几何观点看就是寻求函数 f 在曲面S 上的极值, 更精确些说, 就是考察函数 f 在曲面 S 上的限制 $f|_S$ 并寻求函数 $f|_S$ 的极值.

当然, 这时局部极值点的概念仍保留原来的含义, 即点 $x_0 \in S$ 是函数 f 在 S 上的局部极值点, 或简短地说, $x_0 \in S$ 是函数 $f|_S$ 的局部极值点, 如果存在点 x_0 在集合 $S \subset \mathbb{R}^n$ 中的邻域 $U_S(x_0) = S \cap U(x_0)$(其中 $U(x_0)$ 是 x_0 在 $\mathbb{R}^n$ 中的邻域), 对其中任一点 x, 有不等式 $f(x) \geqslant f(x_0)$(此时称 x_0 *为局部极小值点*) 或 $f(x) \leqslant f(x_0)$(此时称 x_0 *为局部极大值点*). 如果对任一 $x \in U_S(x_0) \setminus x_0$, 上面不等式为严格不等式, 则极值点像以前一样称为*严格局部极值点*.

b. 条件极值的必要条件

定理 1 *设 $f: D \to \mathbb{R}$ 是定义在开集 $D \subset \mathbb{R}^n$ 上且属于 $C^{(1)}(D;\mathbb{R})$ 类的函数, S 是 D 中的光滑曲面且 $x_0 \in S$ 是 f 的非临界点. 如果 x_0 是函数 $f|_S$ 的局部极值点, 则有*

$$\boxed{TS_{x_0} \subset TN_{x_0}.} \tag{26}$$

其中 TS_{x_0} 是曲面S 在 x_0 的切空间, 而 TN_{x_0} 是曲面 $N = \{x \in D | f(x) = f(x_0)\}$(这是与 x_0 有关的函数等高集) 在 x_0 的切空间.

首先我们注意, 从所讨论的求条件极值问题的上下文看, 条件 "x_0 是函数 f 的非临界点" 不是本质的限制. 事实上, 如果 $x_0 \in S$ 已经是函数 $f: D \to \mathbb{R}$ 的临界点或极值点, 那么, 显然它也是函数 $f|_S$ 的可疑点或极值点. 这样一来, 在所考察的问题中真正需要研究的新问题是, 函数 $f|_S$ 可能有那样的临界点和极值点, 它们不是函数 f 的临界点和极值点.

◀ 任取向量 $\xi \in TS_{x_0}$ 与 S 上那样的光滑道路 $x = x(t)$, 它满足当 $t = 0$ 时通过点 x_0 且向量 ξ 就是它在 $t = 0$ 时的速度向量, 即

$$\frac{dx}{dt}(0) = \xi. \tag{27}$$

如果 x_0 是函数 $f|_S$ 的极值点, 则光滑函数 $f(x(t))$ 当 $t = 0$ 时应当有极值, 由极值必要条件知它在 $t = 0$ 时的导数应为零, 即应满足条件

$$f'(x_0) \cdot \xi = 0, \tag{28}$$

其中

$$f'(x_0) = \left(\frac{\partial f}{\partial x^1}, \cdots, \frac{\partial f}{\partial x^n}\right)(x_0), \quad \xi = (\xi^1, \cdots, \xi^n).$$

因为 x_0 是函数 f 的非临界点, 条件 (28) 等价于 $\xi \in TN_{x_0}$. 这是因为关系式 (28) 正是切空间 TN_{x_0} 的方程.

这就证明了 $TS_{x_0} \subset TN_{x_0}$. ▶

如果曲面S 在点 x_0 邻域内用方程组 (25) 表示, 那么, 空间 TS_{x_0} 正像我们知道的那样, 用下面的线性方程组表示

$$\begin{cases} \dfrac{\partial F^1}{\partial x^1}(x_0)\xi^1 + \cdots + \dfrac{\partial F^1}{\partial x^n}(x_0)\xi^n = 0, \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \dfrac{\partial F^m}{\partial x^1}(x_0)\xi^1 + \cdots + \dfrac{\partial F^m}{\partial x^n}(x_0)\xi^n = 0. \end{cases} \tag{29}$$

空间 TN_{x_0} 用下面的方程表示

$$\frac{\partial f}{\partial x^1}(x_0)\xi^1 + \cdots + \frac{\partial f}{\partial x^n}(x_0)\xi^n = 0, \tag{30}$$

又由于方程组 (29) 的所有解是方程 (30) 的解, 所以最后的方程是方程组 (29) 的推论.

从这些考虑推出, 与关系式 $TS_{x_0} \subset TN_{x_0}$ 等价的解析写法是: 向量 $\operatorname{grad} f(x_0)$ 是向量 $\operatorname{grad} F^i(x_0)(i = 1, \cdots, m)$ 的线性组合, 即

$$\operatorname{grad} f(x) = \sum_{i=1}^{m} \lambda_i \operatorname{grad} F^i(x_0). \tag{31}$$

考虑到这个公式是函数 (24) 在约束条件 (25) 下极值的必要条件. 拉格朗日建议在寻求条件极值时利用下面的 $n+m$ 个变量 $(x,\lambda)=(x^1,\cdots,x^n,\lambda_1,\cdots,\lambda_m)$ 的辅助函数

$$L(x,\lambda)=f(x)-\sum_{i=1}^m \lambda_i F^i(x). \tag{32}$$

这个函数称为**拉格朗日函数**, 所用的方法称为**拉格朗日乘数法**.

函数 (32) 的方便之处在于, 它作为 $n+m$ 个变量 $x^1,\cdots,x^n,\lambda_1,\cdots,\lambda_m$ 的函数的极值的必要条件完全与条件 (31) 与 (25) 一致.

事实上, 有

$$\begin{cases} \dfrac{\partial L}{\partial x^j}(x,\lambda)=\dfrac{\partial f}{\partial x^j}(x)-\displaystyle\sum_{i=1}^m \lambda_i \dfrac{\partial F^i}{\partial x^j}(x)=0 & (j=1,\cdots,n) \\ \dfrac{\partial L}{\partial \lambda^i}(x,\lambda)=F^i(x)=0 & (i=1,\cdots,m). \end{cases} \tag{33}$$

这样, 当寻求函数 (24) 在条件 (25) 下的极值时, 可以写出含待定因子的拉格朗日函数 (32), 而要求的就是它的临界点. 如果能从方程组 (33) 求出临界点 $x_0=(x_0^1,\cdots,x_0^n)$ 而不求 $\lambda=(\lambda_1,\cdots,\lambda_m)$, 那么, 从求解原来问题的观点看, 这正是应当做的.

从关系式 (31) 可以看出, 只要向量 $\operatorname{grad} F^i(x_0)(i=1,\cdots,m)$ 线性无关, 因子 $\lambda_i(i=1,\cdots,m)$ 是唯一确定的. 这些向量的无关性等价于方程组 (29) 的秩为 m, 也就是说, 这个方程组的所有方程是独立的 (它们中的任何一个都不能由其他方程推出).

这个要求通常是能满足的, 因为可以认为 (25) 式所有关系式是无关的, 且函数组 $F^1,\cdots,F^m$ 在任一点 $x\in S$ 的秩都是 m.

拉格朗日函数常常写成下面形式

$$L(x,\lambda)=f(x)+\sum_{i=1}^m \lambda_i F^i(x),$$

这个形式不同于以前形式仅仅是作了一个非本质的替换, 即用 λ_i 替换 $-\lambda_i$①.

例 9 求对称二次型

$$f(x)=\sum_{i,j=1}^n a_{ij}x^ix^j \quad (a_{ij}=a_{ji}) \tag{34}$$

在球面

$$F(x)=\sum_{i=1}^n (x^i)^2-1=0 \tag{35}$$

上的极值.

①关于条件极值的必要条件也可参看第十章 (第 II 卷) §7 习题 6.

写出此问题的拉格朗日函数

$$L(x,\lambda)=\sum_{i,j=1}^{n}a_{ij}x^ix^j-\lambda\left(\sum_{i=1}^{n}(x^i)^2-1\right).$$

考虑到 $a_{ij}=a_{ji}$, 求出函数 $L(x,\lambda)$ 极值的必要条件

$$\begin{cases}\dfrac{\partial L}{\partial x^i}(x,\lambda)=2\left(\displaystyle\sum_{j=1}^{n}a_{ij}x^j-\lambda x^i\right)=0\quad(i=1,\cdots,n),\\ \dfrac{\partial L}{\partial\lambda}(x,\lambda)=\left(\displaystyle\sum_{i=1}^{n}(x^i)^2-1\right)=0.\end{cases}\tag{36}$$

把第一个方程组中的每一个方程分别乘以 $x^i(i=1,\cdots,n)$ 然后相加, 并注意到第二个方程, 就得到极值点应满足等式

$$\sum_{i,j=1}^{n}a_{ij}x^ix^j-\lambda=0.\tag{37}$$

方程组 (36) 除去最后一个方程可以改写为

$$\sum_{j=1}^{n}a_{ij}x^j=\lambda x^i\quad(i=1,\cdots,n).\tag{38}$$

由此得到, λ 是由矩阵 (a_{ij}) 确定的线性变换 A 的特征值, 而 $x=(x^1,\cdots,x^n)$ 是这个变换对应于特征值 λ 的特征向量.

因为连续函数 (34) 在紧集 $S=\left\{x\in\mathbb{R}^n\middle|\sum_{i=1}^{n}(x^i)^2=1\right\}$ 上应当在某点具有最大值, 因此方程组 (36) 亦即方程组 (38) 应当有解, 这样我们顺便得到: 任何一个实对称矩阵 (a_{ij}) 至少有一个实特征值. 这是大家从线性代数中已经知道的结果, 这个结果在证明存在由对称算子的特征向量构成的基底时是基本的.

为了指出特征值 λ 的几何意义, 我们发现, 如果 $\lambda>0$, 作坐标变换 $t^i=\dfrac{1}{\sqrt{\lambda}}x^i$, 则 (37) 化作

$$\sum_{i,j=1}^{n}a_{ij}t^it^j=1,\tag{39}$$

而 (35) 化作

$$\sum_{i=1}^{n}(t^i)^2=\frac{1}{\lambda}.\tag{40}$$

但是, $\sqrt{\sum_{i=1}^{n}(t^i)^2}$ 是二次式 (39) 的点 $t=(t^1,\cdots,t^n)$ 到坐标原点的距离. 这样, 如果关系式 (39) 给出一个椭球面, 那么, 特征值 λ 的倒数 $\dfrac{1}{\lambda}$ 是这个椭球的一个半轴之长的平方.

这种观察是有益的, 特别地, 它可以证明条件极值的必要条件 (36) 不是充分的: 要知道, 例如在 $\mathbb{R}^3$ 中的椭球, 除最长与最短半轴外还可能有介于中间的半轴, 在这种中间半轴端点的任意邻域内既含有离坐标原点较近的点, 也含有离坐标原点较远的点 (和该半轴端点至原点的距离相比较而言). 通过中间半轴与最短半轴或最长半轴分别作椭球的两个截面得到两个椭圆, 如果考虑这些椭圆, 上述结论将变得非常显然了. 在第一种情况, 中间半轴是椭圆的两个半轴中的最大者; 在第二种情况, 它是最小者.

应当补充说明, 如果 $1/\sqrt{\lambda}$ 是这个中间半轴的长, 那么, 从椭球的标准方程可以看出, 数 λ 显然是变换 A 的特征值, 因此, 表示函数 $f|_S$ 极值必要条件的方程组 (36) 实际上有解, 但它不是这个函数的极值点.

定理 1 得到的结果 (条件极值的必要条件) 可用图 63(a),(b) 表示.

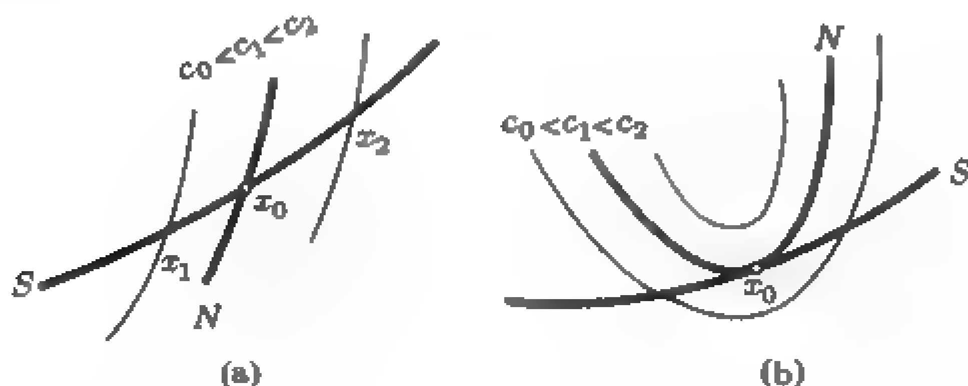


图 63

图 63(a) 说明, 为什么当曲面S 与曲面$N = \{x \in \mathbb{R}^n | f(x) = f(x_0) = c_0\}$ 在点 x_0 不相切时, 曲面S 上的点 x_0 不可能是函数 $f|_S$ 的极值点. 这时假设了 grad $f(x_0) \neq 0$, 这个条件保证了在点 x_0 的邻域内既有函数 f 较高的 c_2- 等高集的点, 也有这个函数较低的 c_1- 等高集的点.

因为光滑曲面S 与曲面N(即光滑函数 f 的 c_0- 等高集) 相交, 那么, 在点 x_0 的邻域内, S 既与函数 f 的较高等高集相交, 也与它的较低等高集相交. 而这意味着, 点 x_0 不可能是函数 $f|_S$ 的极值点.

图 63(b) 说明, 为什么当 N 与 S 在点 x_0 相切时, 这个点可能是极值点. 图中点 x_0 是函数 $f|_S$ 的局部极大值点.

就是这样一些思考, 使我们能画出图来, 而图的解析形式能证明条件极值的必要条件不是充分的.

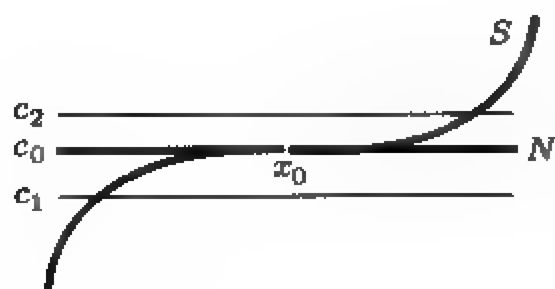


图 64

事实上, 例如, 对应图 64 设

$$f(x,y)=y,$$
$$F(x,y)=x^3-y=0.$$

这时, 曲线 $S\subset\mathbb{R}^2$ 显然用方程 $y=x^3$ 给出, 量 y 在点 (0,0) 不具极值, 虽然这条曲线与函数 f 在这点的水平集 $f(x,y)=0$ 相切. 注意, $\operatorname{grad} f(0,0)=(0,1)\neq 0$.

显然, 这个例子本质上就是我们在研究古典的函数内部极值点的充分条件与必要条件之间的差别时所使用的那个例子.

c. 条件极值的充分条件 现在我们证明下面的条件极值存在或不存在的充分条件.

定理 2 *设 $f: D\to\mathbb{R}$ 是定义在开集 $D\subset\mathbb{R}^n$ 上且属于 $C^{(2)}(D;\mathbb{R})$ 类的函数, S 是用方程组 (25) 表示的 D 中的曲面, 其中 $F^i\in C^{(2)}(D;\mathbb{R})(i=1,\cdots,m)$ 且函数组 $\{F^1,\cdots,F^m\}$ 在区域 D 中任一点的秩等于 m.*

设拉格朗日函数

$$L(x)=L(x,\lambda)=f(x^1,\cdots,x^n)-\sum_{i=1}^{m}\lambda_i F^i(x^1,\cdots,x^m)$$

中的参数 $\lambda_1,\cdots,\lambda_m$ 已根据函数 $f|_S$ 在点 $x_0\in S$ 取极值的必要条件 (31) 选定[①].

如果二次型

$$\frac{\partial^2 L}{\partial x^i\partial x^j}(x_0)\xi^i\xi^j \tag{41}$$

对任意的向量 $\xi\in TS_{x_0}$ 有确定的符号, 则点 x_0 是函数 $f|_S$的极值点.

如果 (41) 式在 TS_{x_0} 上是正定的, 则 x_0 是函数 $f|_S$ 的严格局部极小值点; 如果 (41) 式在 TS_{x_0} 上是负定的, 则 x_0 是函数 $f|_S$ 的严格局部极大值点.

如果 (41) 式在 TS_{x_0} 上的值具有不同的符号, 则 x_0 不是函数 $f|_S$ 的极值点.

◀ 首先注意到, 对于 $x\in S, L(x)\equiv f(x)$, 因此, 如果我们证明了点 $x_0\in S$ 是函数 $L|_S$ 的极值点, 也就同时证明了它是函数 $f|_S$ 的极值点.

由于函数 $f|_S$ 在点 $x_0\in S$ 满足极值的必要条件 (31), 因此在这点有 $\operatorname{grad} L(x_0)=0$, 这意味着函数 $L(x)$ 在点 $x_0=(x_0^1,\cdots,x_0^n)$ 邻域内的泰勒展开式为

$$L(x)-L(x_0)=\frac{1}{2!}\sum_{i,j}\frac{\partial^2 L}{\partial x^i\partial x^j}(x_0)(x^i-x_0^i)(x^j-x_0^j)+o(\|x-x_0\|^2), 当 x\to x_0. \tag{42}$$

现在, 我们想起引进定义 2 的原因, 发现 k 维光滑曲面 (此时 $k=n-m$) 在局部范围 (例如在点 $x_0\in S$ 的邻域) 内可以写成参数形式.

① λ 确定后, 从 $L(x,\lambda)$ 得到只依赖于 x 的函数, 从而可用 $L(x)$ 表示它.

换句话说, 存在光滑映射

$$\mathbb{R}^k \ni (t^1, \cdots, t^k) = t \mapsto x = (x^1, \cdots, x^n) \in \mathbb{R}^n$$

(像以前一样, 我们将它写为 $x = x(t)$), 它把点 $0 = (0, \cdots, 0) \in \mathbb{R}^k$ 的邻域双方单值地映成点 $x_0 \in S$ 的某个邻域且 $x_0 = x(0)$.

注意到关系式

$$x(t) - x(0) = x'(0) + o(\|t\|), \quad \text{当} \, t \to 0$$

(它表现了映射 $x = x(t)$ 在点 $t = 0$ 的可微性) 等价于 n 个坐标等式

$$x^i(t) - x^i(0) = \frac{\partial x^i}{\partial t^\alpha}(0)t^\alpha + o(\|t\|) \quad (i = 1, \cdots, n) \tag{43}$$

上式右边是关于指标 α 从 1 到 k 求和.

由此推出

$$|x^i(t) - x^i(0)| = o(\|t\|), \quad \text{当} \, t \to 0,$$

即

$$\|x(t) - x(0)\|_{\mathbb{R}^n} = o(\|t\|_{\mathbb{R}^k}), \quad \text{当} \, t \to 0, \tag{44}$$

利用关系式 (43),(44), 从等式 (42) 得到

$$L(x(t)) - L(x(0)) = \frac{1}{2!}\partial_{ij}L(x_0)\partial_\alpha x^i(0)\partial_\beta x^j(0)t^\alpha t^\beta + o(\|t\|^2), \quad \text{当} \, t \to 0. \tag{42'}$$

由此, 在二次型

$$\partial_{ij}L(x_0)\partial_\alpha x^i(0)\partial_\beta x^j(0)t^\alpha t^\beta \tag{45}$$

定号条件下, 得到函数 $L(x(t))$ 在 $t = 0$ 有极值; 如果 (45) 式有不同的符号, 则 $L(x(t))$ 在 $t = 0$ 没有极值. 但是, 因为映射 $t \mapsto x(t)$ 把点 $0 \in \mathbb{R}^k$ 某个邻域变到曲面 S 上点 $x(0) = x_0 \in S$ 的邻域, 可以断定, 函数 $L|_S$ 在点 x_0 或者有极值, 当函数 $L(x(t))$ 在 $t = 0$ 有极值; 或者像函数 $L(x(t))$ 一样没有极值.

于是, 剩下来的是要验证: 对任何向量 $\xi \in TS_{x_0}$, 表达式 (41) 与 (45) 只是同一个量的不同表示.

事实上, 设

$$\xi = x'(0)t,$$

由此知向量 ξ 在点 x_0 与 S 相切, 记

$$\xi = (\xi^1, \cdots, \xi^n), \quad x(t) = (x^1, \cdots, x^n)(t), \quad t = (t^1, \cdots, t^k),$$

则

$$\xi^j = \partial_\beta x^j(0)t^\beta \quad (j = 1, \cdots, n),$$

由此推出 (41) 与 (45) 是一样的. ►

我们发现利用定理 2 的实际困难是: 在向量 $\xi = (\xi^1, \cdots, \xi^n) \in TS_{x_0}$ 的坐标中只有 $k = n - m$ 个是无关的. 这是因为向量 ξ 的坐标应当满足确定切空间 TS_{x_0} 的方程组 (29). 因此, 在这种情况, 对于公式 (41) 直接应用西尔维斯特准则, 一般来说, 是什么也得不到的: 二次型 (41) 在 $T\mathbb{R}^n_{x_0}$ 上可能是不定的, 但在 TS_{x_0} 上确是定的. 如果从 (29) 中把向量 ξ 的 m 个坐标用其余的 k 个独立坐标表示出来, 并将所得的线性形式代入 (41), 那么, 我们就得到具有 k 个独立变量的二次型, 这个二次型的确定性问题已经可以利用西尔维斯特准则来研究了.

现在举出一些最简单的例子说明上边所讲的内容.

例 10 设在空间 $\mathbb{R}^3$ 中用坐标 x、y、z 给出函数

$$f(x,y,z) = x^2 - y^2 + z^2.$$

寻求这个函数在用方程

$$F(x,y,z) = 2x - y - 3 = 0$$

给出的平面 S 上的极值.

写出拉格朗日函数

$$L(x,y,z) = (x^2 - y^2 + z^2) - \lambda(2x - y - 3)$$

和极值的必要条件

$$\begin{cases} \dfrac{\partial L}{\partial x} = 2x - 2\lambda = 0, \\ \dfrac{\partial L}{\partial y} = -2y + \lambda = 0, \\ \dfrac{\partial L}{\partial z} = 2z = 0, \\ \dfrac{\partial L}{\partial \lambda} = -(2x - y - 3) = 0, \end{cases}$$

求出可疑点 $p = (2,1,0)$.

其次, 求出公式 (41):

$$\frac{1}{2}\partial_{ij}L\xi^i\xi^j = (\xi^1)^2 - (\xi^2)^2 + (\xi^3)^2, \tag{46}$$

我们看到, 在这里, 参数 λ 没有在二次型中出现, 所以不必去计算它.

现在写出条件 $\xi \in TS_p$:

$$2\xi^1 - \xi^2 = 0, \tag{47}$$

从这个等式得到 $\xi^2 = 2\xi^1$, 代入 (46), 使它化成

$$-3(\xi^1)^2 + (\xi^3)^2.$$

其中 ξ^1 与 ξ^3 是独立变量.

最后这个二次型显然可取不同的符号, 因此函数 $f|_S$ 在点 $p \in S$ 没有极值.

例 11 在例 10 中用 $\mathbb{R}^2$ 替换 $\mathbb{R}^3$, 函数 f 换成

$$f(x, y) = x^2 - y^2,$$

且保存条件

$$2x - y - 3 = 0,$$

现在它给出平面 $\mathbb{R}^2$ 上的一条直线 S.

求出可疑点 $p = (2, 1)$.

代替 (46) 得到二次型

$$(\xi^1)^2 - (\xi^2)^2, \tag{48}$$

以及 ξ^1 和 ξ^2 之间的关系式 (47). 这样一来, 在 TS_p 上 (48) 式可写成

$$-3(\xi^1)^2,$$

也就是说, 它是负定的. 由此断定, 点 $p = (2, 1)$ 是函数 $f|_S$ 的局部极大值点.

下面这些简单例子在许多方面都是非常有教益的, 在这些例子中可以清楚地看到极值必要条件, 充分条件的作用机制, 也包括参数的作用和拉格朗日函数的非形式的作用.

例 12 在具笛卡儿坐标 (x, y) 的平面上给出函数

$$f(x, y) = x^2 + y^2$$

我们来求这个函数在由

$$F(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0 \quad (0 < a < b)$$

定义的椭圆周 S 上的极值.

从几何上考虑, 显然有 $\min f|_S = a^2, \max f|_S = b^2$. 下面, 我们利用定理 1 和定理 2 导出这个结果.

写出拉格朗日函数

$$L(x, y, \lambda) = (x^2 + y^2) - \lambda\left(\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1\right),$$

并求解方程 $dL=0$, 亦即方程组 $\dfrac{\partial L}{\partial x}=\dfrac{\partial L}{\partial y}=\dfrac{\partial L}{\partial \lambda}=0$, 找出它的解:

$$(x,y,\lambda)=(\pm a,0,a^2),(0,\pm b,b^2).$$

按照定理 2, 现在写出二次形式 $\frac{1}{2}d^2L\xi^2$ 并研究它, 它是拉格朗日函数在相应点邻域中的泰勒展开式中的第二项:

$$\frac{1}{2}d^2L\xi^2=\left(1-\frac{\lambda}{a^2}\right)(\xi^1)^2+\left(1-\frac{\lambda}{b^2}\right)(\xi^2)^2.$$

在椭圆周 S 的点 $(\pm a,0)$ 处, 切向量$\xi=(\xi^1,\xi^2)$ 为 $(0,\xi^2)$, 而当 $\lambda=a^2$ 时二次型为

$$\left(1-\frac{a^2}{b^2}\right)(\xi^2)^2.$$

注意到条件 $0<a<b$, 得知, 这个二次型是正定的, 因此, 在点 $(\pm a,0)\in S$ 函数 $f|_S$ 有严格局部极小值 (而且, 也是整体最小值), 即 $\min f|_S=a^2$.

类似地, 求出二次型

$$\left(1-\frac{b^2}{a^2}\right)(\xi^1)^2,$$

它对应于点 $(0,\pm b)\in S$, 并得到 $\max f|_S=b^2$.

注　注意比较拉格朗日函数和函数 f 的作用. 在点 $(\pm a,0),(0,\pm b)$ 函数 f (与函数 L 一样) 沿相应切向量的微分都是零, 而二次型 $\frac{1}{2}d^2f\xi^2=(\xi^1)^2+(\xi^2)^2$, 无论在这些点的哪一个点算出来都是这样, 然而, 函数 $f|_S$ 在点 $(\pm a,0)$ 有严格最小, 而在点 $(0,\pm b)$ 有严格最大.

为了弄明白, 这到底是怎么回事, 请再一次细细考察定理 2 的证明并且试着把 (42) 中的 L 换成 f, 得到关系式 (42′). 你会发现, 这时将出现一个包含 $x''(0)$ 的附加项. 有这样一项的原因是, 函数 f 的微分与 dL 不同, 它在相应的点处不恒等于零, 虽然它沿切向量 (形如 $x'(0)$) 的值等于零.

例 13　求函数

$$f(x,y,z)=x^2+y^2+z^2$$

在由关系式

$$F(x,y,z)=\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}-1=0\quad(0<a<b<c)$$

定义的椭球面 S 上的极值.

写出拉格朗日函数

$$L(x,y,z,\lambda)=(x^2+y^2+z^2)-\lambda\left(\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}-1\right),$$

根据极值必要条件求出方程 $dL=0$, 亦即 $\frac{\partial L}{\partial x}=\frac{\partial L}{\partial y}=\frac{\partial L}{\partial z}=\frac{\partial L}{\partial \lambda}=0$ 的解:

$$(x,y,z,\lambda)=(\pm a,0,0,a^2),(0,\pm b,0,b^2),(0,0,\pm c,c^2).$$

二次型

$$\frac{1}{2}d^2L\xi^2=\left(1-\frac{\lambda}{a^2}\right)(\xi^1)^2+\left(1-\frac{\lambda}{b^2}\right)(\xi^2)^2+\left(1-\frac{\lambda}{c^2}\right)(\xi^3)^2$$

对每个可疑点在相应的切平面上分别有

$$\left(1-\frac{a^2}{b^2}\right)(\xi^2)^2+\left(1-\frac{a^2}{c^2}\right)(\xi^3)^2, \tag{a}$$

$$\left(1-\frac{b^2}{a^2}\right)(\xi^1)^2+\left(1-\frac{b^2}{c^2}\right)(\xi^3)^2, \tag{b}$$

$$\left(1-\frac{c^2}{a^2}\right)(\xi^1)^2+\left(1-\frac{c^2}{b^2}\right)(\xi^2)^2. \tag{c}$$

因为 $0<a<b<c$, 根据给出条件极值存在性和不存在性的充分条件的定理 2, 在 (a) 和 (c) 两种情形, 分别有 $\min f|_S=a^2$ 和 $\max f|_S=c^2$, 而在点 $(0,\pm b,0)\in S$, 相应于情形 (b), 函数 $f|_S$ 没有极值. 这与在讨论条件极值的必要条件时所说的几何上的一些明显的想法是完全一致的.

本节中分析概念和几何概念所涉及的一些进一步的, 有时是极有益的内容, 其中包括条件极值问题本身的物理解释, 也包括极值必要条件 (31) 作为力在平衡点处的分解和拉格朗日乘数作为理想约束的反作用量这样一些物理解释, 在下面的练习中作了介绍.

练　习

1. 道路与曲面

a) 设 $f:I\to\mathbb{R}^2$ 是定义在区间 $I\subset\mathbb{R}$ 上且属于 $C^{(1)}(I;\mathbb{R})$ 类的函数. 把这个映射看作 $\mathbb{R}^2$ 中的道路. 举例证明它的承载子 $f(I)$ 可能不是 $\mathbb{R}^2$ 中的子流形, 而它在 $\mathbb{R}^3=\mathbb{R}^1\times\mathbb{R}^2$ 中的图像却总是 $\mathbb{R}^3$ 中的一维子流形, 这个子流形在 $\mathbb{R}^2$ 上的投影正是上面所指出的承载子 $f(I)$.

b) 当 I 是 $\mathbb{R}^k$ 中的区间, $f\in C^{(1)}(I;\mathbb{R}^n)$ 时, 证明: 映射 $f:I\to\mathbb{R}^n$ 的图像是 $\mathbb{R}^k\times\mathbb{R}^n$ 中的 k 维光滑曲面, 这个曲面在子空间 $\mathbb{R}^n$ 上的投影正是 $f(I)$.

c) 证明: 如果 $f_1:I_1\to S, f_2:I_2\to S$ 是同一个 k 维曲面 $S\subset\mathbb{R}^n$ 的两个光滑参数式, 并且无论 f_1 在 I_1 上或者 f_2 在 I_2 上都没有临界点, 则映射 $f_1^{-1}\circ f_2:I_2\to I_1$ 与映射 $f_2^{-1}\circ f_1:I_1\to I_2$ 是光滑的.

2. $\mathbb{R}^n$ 中的球面

a) 利用 $\mathbb{R}^3$ 中的极坐标 (参看上节公式 (5)) 令 $\rho=1$ 就得到球面 $S^2=\{x\in\mathbb{R}^3 \mid \|x\|=1\}$ 上的曲线坐标 (φ,ψ), 指出这个曲线坐标的一个最大有效范围.

b) 当 $S^{m-1}=\{x\in\mathbb{R}^m \mid \|x\|=1\}$ 是 $\mathbb{R}^m$ 中的 $(m-1)$ 维球面, $(\varphi_1,\cdots,\varphi_{m-1})$ 是由 $\mathbb{R}^m$ 中的极坐标 (参看上节公式 (6)) 令 $\rho=1$ 得到的球面曲线坐标时, 回答问题 a).

c) 能否用一个坐标系 $(t^1,\cdots,t^k)$, 即一个区域 $G\subset\mathbb{R}^k$ 上的微分同胚 $f:G\to\mathbb{R}^{k+1}$ 给出球面 $S^k\subset\mathbb{R}^{k+1}$?

d) 在地表图册中至少应当有几张地图?

e) 球面 $S^2\subset\mathbb{R}^3$ 上两点之间的距离指的是联结这两点的球面 S^2 上最短曲线的长, 这样的曲线是相应的大圆周的弧. 能否存在这样局部平的球面地图, 使球面上点之间的距离与在地图上表示它们的点之间的距离成比例 (具有同一比例系数).

f) 称相交于一点的两条曲线 (属于或不属于球面) 在这点的切线之间的夹角为这两条曲线在这点的夹角.

证明: 存在球面的局部平地图, 在这个地图中, 地面上曲线之间的交角和地图上对应的曲线之间的夹角是相同的 (参看图 65, 图中表示的是所谓的球面投影).

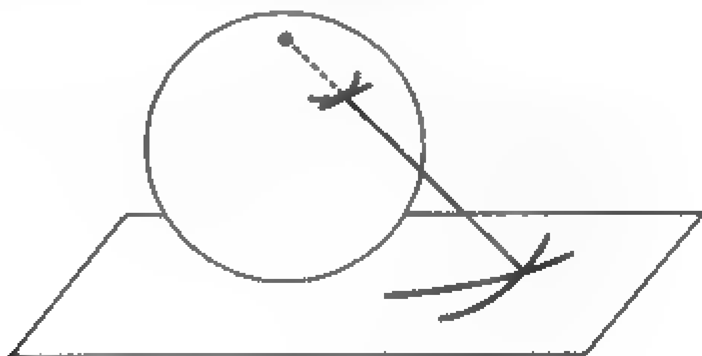


图 65

3. 切空间

a) 通过直接计算验证 k 维光滑曲面 $S\subset\mathbb{R}^n$ 在点 $x_0\in S$ 的切流形 TS_{x_0} 不依赖于 $\mathbb{R}^n$ 中坐标系的选择.

b) 证明: 如果从区域 $D\subset\mathbb{R}^n$ 到区域 $D'\subset\mathbb{R}^n$ 的微分同胚 $f:D\to D'$ 把光滑曲面 $S\subset D$ 映成光滑曲面 $S'\subset D'$, 把点 $x_0\in S$ 映成点 $x_0'\in S'$, 则 f 在点 $x_0\in D$ 的切变换 (线性变换) $f'(x_0):\mathbb{R}^n\to\mathbb{R}^n$ 是向量空间 TS_{x_0} 到向量空间 TS'_{x_0} 的同构变换.

c) 如果上面问题中的映射 $f:D\to D'$ 属于 $C^{(1)}(D;D')$ 类且 $f(S)\subset S'$, 则 $f'(TS_{x_0})\subset TS'_{x_0'}$.

d) 证明: 从 k 维光滑曲面 $S\subset\mathbb{R}^n$ 到它在点 $x_0\in S$ 处的 k 维切平面 TS_{x_0} 上的正交投影映射, 在切点 x_0 的某个邻域内是双方单值的.

e) 在上面问题的条件中, 设 $\xi\in TS_{x_0}$ 且 $\|\xi\|=1$.

可以借助 $\mathbb{R}^n$ 中属于 TS_{x_0} 的直线方程 $x-x_0=\xi t$ 用序对 (t,ξ) 刻画每个点 $x\in TS_{x_0}\setminus x_0$. 这实质上是 TS_{x_0} 中的极坐标. 证明: 在曲面上点 x_0 的邻域内, 直线 $x-x_0=\xi t$ 与曲面上仅在 x_0 相交的一些光滑曲线相对应. 验证: 在这些曲线上保留 t 为参数, 我们得到一些道路, 沿着它们的速度当 $t=0$ 时与确定直线 $x-x_0=\xi t$ 的向量 $\xi\in TS_{x_0}$ 重合.

这样, 在曲面 S 上点 $x_0 \in S$ 的某邻域内, 可把序对 (t,ξ)(其中 $\xi \in TS_{x_0}, \|\xi\|=1, t$ 是 $\mathbb{R}$ 中原点的某个邻域 $U(0)$ 中的实数) 作为 S 的点的曲线坐标, 它类似于极坐标.

4. 设函数 $F \in C^{(1)}(\mathbb{R}^n;\mathbb{R})$ 没有临界点, 且用方程

$$F(x^1,\cdots,x^n)=0$$

给出 $\mathbb{R}^n$ 中的紧曲面 S(即 S 是 $\mathbb{R}^n$ 中的紧子集). 对任一点 $x \in S$, 找到向量 $\eta(x) = \operatorname{grad} F(x)$, 它在点 x 垂直于 S. 如果每个点 $x \in S$, 都以自己的速度 $\eta(x)$ 匀速移动, 则产生出一个依赖于时间 t 的映射 $S \ni x \mapsto x+\eta(x)t \in \mathbb{R}^n$.

a) 证明: 当 t 充分接近零时, 这个映射是双射, 而且, 对每个这样的 t, 它把 S 映成光滑曲面 $\tilde{S}_t$.

b) 设 E 是 $\mathbb{R}^n$ 中的集合, 称 $\mathbb{R}^n$ 中一切到 E 的距离小于 δ 的点的集合为E 的 δ 邻域. 证明: 当 t 接近零时, 方程

$$F(x^1,\cdots,x^n)=t$$

给出紧曲面 $S_t \subset \mathbb{R}^n$, 并证明: 曲面 $\tilde{S}_t$ 含在 S_t 的 $\delta(t)$ 邻域之中, 其中当 $t \to 0$ 时, $\delta(t)=o(t)$.

c) 对曲面 $S=S_0$ 的每个点 x 取单位法向量 $n(x)=\dfrac{\eta(x)}{\|\eta(x)\|}$, 考察新映射 $S \ni x \mapsto x+n(x)t \in \mathbb{R}^n$. 证明: 对于所有充分接近于零的 t, 这个映射把 S 双方单值地映成光滑曲面 $\tilde{\tilde{S}}_t$ 且当 $t_1 \neq t_2$ 时, 有 $\tilde{\tilde{S}}_{t_1} \cap \tilde{\tilde{S}}_{t_2} = \varnothing$.

d) 根据上题结果, 证明: 存在 $\delta > 0$, 使得曲面 S 的 δ 邻域中的点与序对 (t,x)(其中 $t \in]-\delta,\delta[\subset \mathbb{R}, x \in S$) 之间有相互单值的对应; 如果 $(t^1,\cdots,t^k)$ 是在点 x_0 的邻域 $U_S(x_0)$ 内曲面 S 上的局部坐标, 则 $(t,t^1,\cdots,t^k)$ 可以作为点 $x_0 \in \mathbb{R}^n$ 的某个空间邻域 $U(x_0)$ 内的局部坐标.

e) 证明: 当 $|t|<\delta$ 时, 点 $x \in S$ 是曲面 S 上距点 $x+n(x)t \in \mathbb{R}^n$ 最近的点. 这样, 曲面 $\tilde{\tilde{S}}_t$ 当 $|t|<\delta$ 时是空间 $\mathbb{R}^n$ 中离开曲面 S 的距离等于 $|t|$ 的点的几何轨迹.

5. a) 设 $d_p: S \to \mathbb{R}$ 是 k 维光滑曲面 $S \subset \mathbb{R}^n$ 上由 $d_p(x)=\|p-x\|^2$ 定义的函数, 其中 p 是 $\mathbb{R}^n$ 中确定的点, $x \in S$ 而 $\|p-x\|$ 是 $\mathbb{R}^n$ 中点 p 与点 x 的距离. 证明: 在函数 $d_p(x)$ 的极值点处向量 $p-x$ 正交于曲面 S.

b) 证明: 在点 $q \in S$ 处垂直于曲面 S 的任一直线上至多有 k 个点 p 使得点 q 是函数 $d_p(x)$ 的退化临界点 (即函数在此点的黑塞矩阵为零阵).

c) 证明: 如果曲线 $S(k=1)$ 在平面 $\mathbb{R}^2(n=2)$ 上, 而点 p 使点 $q \in S$ 是函数 $d_p(x)$ 的退化临界点, 则点 p 与 S 在点 $q \in S$ 的曲率中心重合.

6. 在笛卡儿坐标 x,y 的平面 $\mathbb{R}^2$ 上画出函数 $f(x,y)=xy$ 的等高集和曲线 $S=\{(x,y)\in \mathbb{R}^2 | x^2+y^2=1\}$. 利用所得到的图形对函数 $f|_S$ 的极值问题进行充分的研究.

7. 在笛卡儿坐标 x,y 的平面 $\mathbb{R}^2$ 上定义了下面的 $C^{(\infty)}(\mathbb{R}^2;\mathbb{R})$ 类函数:

$$f(x,y)=x^2-y;$$

$$F(x,y)=\begin{cases} x^2-y+e^{-\frac{1}{x^2}}\sin\dfrac{1}{x}, & x\neq 0,\\ x^2-y, & x=0.\end{cases}$$

a) 画出函数 $f(x,y)$ 的等高集 (等高曲线) 和由关系式 $F(x,y)=0$ 给出的曲线 S.

b) 研究函数 $f|_S$ 的极值.

c) 证明: 二次型 $\partial_{ij}f(x_0)\xi^i\xi^j$ 在 TS_{x_0} 上定型的条件, 与定理 2 中的二次型 $\partial_{ij}L(x_0)\xi^i\xi^j$ 在 TS_{x_0} 上定型的条件不同并非判断函数 $f|_S$ 可疑点 $x_0\in S$ 为极值点的充分条件.

d) 试检验, 点 $x_0=(0,0)$ 是不是函数 f 的临界点, 以及, 能否像 c) 中那样仅借助于泰勒公式的第二项 (即二次项) 研究函数 f 在该点邻域中的性态.

8. 在微分几何中确定主曲率与主方向时, 善于求解一个二次型 $h_{ij}u^iu^j$ 在另外一个正定的二次型 $g_{ij}u^iu^j$ 保持不变的条件下的极值, 常常是很有好处的. 试用类似于本节例 9 的方法解决这个问题.

9. 设 $A=[a_j^i]$ 是 n 阶方阵且 $\sum\limits_{i=1}^n(a_j^i)^2=H_j \quad (j=1,\cdots,n)$, 其中 $H_1,\cdots,H_n$ 是 n 个确定的非负实数.

a) 证明: 仅当矩阵 A 的行向量是 $\mathbb{R}^n$ 中两两正交的向量时, $\det^2 A$ 有极值.

b) 根据等式 $\det^2 A=\det A\cdot\det A^*$, 其中 A^* 是矩阵 A 的转置矩阵, 证明: 在上述条件下, $\max\limits_A \det^2 A=H_1\cdot\cdots\cdot H_n$.

c) 证明: 对任意的矩阵 $[a_j^i]$ 有阿达马不等式

$$\det{}^2(a_j^i)\leqslant\prod_{j=1}^n\left(\sum_{i=1}^n(a_j^i)^2\right).$$

d) 给阿达马不等式以直观的几何解释.

10. a) 描绘本节例 10 中函数 f 的等高集 (等高曲面) 和 S. 用图说明这个例子所得到的结论.

b) 描绘本节例 11 中函数 f 的等高集 (等高曲线) 和直线 S. 用图说明这个例子所得到的结论.

11. 在第五章 §4 的例 6 中, 从费马原理出发得到了光线在两介质分界面与平面情形的斯涅耳折射定律. 试问, 对任意光滑界面情形, 该定律是否仍然有效.

12. a) 在势场力作用下的质点仅在势的临界点(稳定点) 可以处在平衡位置 (也叫静态或稳定态). 同时, 势的严格局部极小对应稳定平衡位置, 而局部极大对应于不稳定位置. 试证这些结果.

b) 试问, 在势场力 (如重力) 作用下且具理想约束 (例如, 点不能离开某光滑曲面, 或珠子不能离开光滑细线, 或小球不能滑出沟槽) 的质点平衡问题归结为怎样的条件极值问题? 约束是理想的 (没有摩擦), 这就是说, 它对质点的作用 (约束的反作用) 只发生在约束的法线方向上.

c) 在这种情况, 展开式 (31), 即条件极值的必要条件和拉格朗日乘数, 有怎样的物理 (力学) 的意义?

顺便指出, 方程组 (25) 中的每个函数都可除以它的模, 显然,(如果它的秩处处等于 m) 这样得到的是一个等价的方程组. 因此, 可以认为, 关系式 (31) 右边的所有的矢量 $\operatorname{grad} F^i(x_0)$ 是相应曲面的单位法向量.

d) 在领悟到物理解释后, 求条件极值的拉格朗日方法本身是否变得自明和自然了呢?

口试试题

分析引论 (数, 函数, 极限)

1. 设沿地球赤道有紧束地球的一个箍. 如果把它增长 1 米, 箍与地面之间将形成一个缝隙. 试问, 这个缝隙是否能让一个蚂蚁通过? 当赤道有这样增长时, 地球半径的绝对增长和相对增长是多少?(地球半径 ≈6400 km)
2. 实数集的完备性 (连续性), 自然数列的无界性和阿基米德原理有什么联系? 为什么任何实数都用有理数以任意精度逼近? 以有理分式 (有理函数) 为例说明阿基米德原理可能失效, 以及在那种数系中自然数序列是有界的而且存在无穷小数.
3. 在单位正方形的四个顶点上各有一只小甲虫, 它们按同一方向以单位速度沿正方形边缘追逐. 试绘出它们的运动轨迹, 每条轨迹的长是多少?(在笛卡儿坐标系和极坐标系中的) 运动规律如何?
4. 画出用迭代法
$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$$
计算 $\sqrt{a}(a > 0)$ 的框图.

   怎样把解方程和求不动点两个问题联系起来? 怎样求 $\sqrt[k]{a}$?
5. 设 $g(x) = f(x) + o(f(x))$(当 $x \to \infty$). 试问, 这时也有 $f(x) = g(x) + o(g(x))$(当 $x \to \infty$) 吗?
6. 用待定系数法 (或其他方法) 求出 $(1+x)^\alpha(\alpha = -1, -\frac{1}{2}, 0, \frac{1}{2}, 1, \frac{3}{2})$ 的幂级数的头几个 (或全部) 系数 (牛顿对这样一些展开式中 x 同次幂的系数进行插值, 提出了对任意 $\alpha \in \mathbb{R}$ 的系数构成规律, 即牛顿二项式)
7. 已知函数 e^x 的幂级数展开式, 试用待定系数法 (或其他方法) 求出函数 $\ln(1+x)$ 的幂级数展开式的头几个 (或全部) 系数.

8. 计算 $\exp A$, 这里 A 是下述矩阵之一:

$$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

9. 为了以 10^{-2} 的精度在区间 $[-3,5]$ 上计算 e^x, 应当将 e^x 作幂级数展开到第几项?

10. 画出下列函数的图像的草图:

a) $\log_{\cos x}\sin x$;　b) $\operatorname{arctg}\dfrac{x^3}{(1-x)(1+x)^2}$

单变量函数的微分学

1. 试证, 如果运动的加速度向量 $a(t)$ 在每一时刻 t 都垂直于速度向量 $v(t)$, 则量 $|v(t)|$ 是常数.

2. 设 (x,t) 和 $(\tilde{x},\tilde{t})$ 是一个运动的点在两个坐标系中相应的坐标. 如果已知从一个坐标系到另一个坐标系的转换公式 $\tilde{x}=\alpha x+\beta t, \tilde{t}=\gamma x+\delta t$, 试求速度变换公式, 即 $v=\dfrac{dx}{dt}$ 和 $\tilde{v}=\dfrac{d\tilde{x}}{d\tilde{t}}$ 的关系式.

3. 函数 $f(x)=x^2\sin\dfrac{1}{x}$ 当 $x\neq 0$, 且 $f(0)=0$, 它在 $\mathbb{R}$ 上可微, 但 f' 在 $x=0$ 间断 (请验证之). 但是, 下边却 "证明" 了: 如果 $f:\mathbb{R}\to\mathbb{R}$ 在 $\mathbb{R}$ 上可微, 则 f' 在每一个点 $a\in\mathbb{R}$ 都连续.

根据拉格朗日定理, 有 $\dfrac{f(x)-f(a)}{x-a}=f'(\xi)$, 这里 ξ 是介于 a 与 x 之间的一个点. 因此, 如果 $x\to a$, 则 $\xi\to a$. 由定义, $\lim\limits_{x\to a}\dfrac{f(x)-f(a)}{x-a}=f'(a)$. 于是, 由这个极限存在就得到, 拉格朗日公式右边的极限也等于 $f'(a)$, 即当 $\xi\to a$ 时有 $f'(\xi)\to f'(a)$. 这样, f' 在点 a 处的连续性 "得证".

请指出这个 "证明" 的错误何在?

4. 设函数 f 在点 x_0 有 $n+1$ 阶导数, 而 $\xi=x_0+\theta_x(x-x_0)$ 是拉格朗日公式中余项 $\dfrac{1}{n!}f^{(n)}(\xi)(x-x_0)^n$ 的中间点, 且 $0<\theta_x<1$, 试证, 如果 $f^{(n+1)}(x_0)\neq 0$, 则当 $x\to x_0$ 时有 $\theta_x\to\dfrac{1}{n+1}$.

5. 试证: 如果数 $a_1,\cdots,a_n,\alpha_1,\cdots,\alpha_n$ 非负且 $\alpha_1+\cdots+\alpha_n=1$, 则成立不等式

$$a_1^{\alpha_1}\cdots a_n^{\alpha_n}\leqslant \alpha_1a_1+\cdots+\alpha_na_n.$$

6. 试证.

$$\lim_{n\to\infty}\left(1+\frac{z}{n}\right)^n=e^x(\cos y+i\sin y)\quad(z=x+iy).$$

由此, 自然认为, $e^{iy}=\cos y+i\sin y$(欧拉公式) 且

$$e^z=e^x\cdot e^{iy}=e^x(\cos y+i\sin y).$$

7. 试求杯子中匀速转动的液体的表面形状.

8. 设 $a>b>0$. 试证: 椭圆 $\dfrac{x^2}{a^2}+\dfrac{y^2}{b^2}=1$ 在它的点 (x_0,y_0) 处的切线方程为 $\dfrac{xx_0}{a^2}+\dfrac{yy_0}{b^2}=1$, 另外, 由位于两焦点 $F_1=(-\sqrt{a^2-b^2},0)$, $F_2=(\sqrt{a^2-b^2},0)$ 之一的光源发出的光线, 必由椭圆反射镜聚到另一个焦点.

9. 一个物体从椭圆形断面冰山顶在重力作用下开始下滑, 设没有预先的驱动, 冰山断面方程为 $x^2+5y^2=1, y\geqslant 0$. 试计算物体运动到它着地的轨道.

积分和多元分析引论

1. 若已知关于和的赫尔德不等式, 闵可夫斯基不等式和詹生不等式, 求关于积分的相应的不等式.

2. 计算积分 $\int_0^1 e^{-x^2}dx$ 至相对误差不超过 10%.

3. 函数 $\mathrm{erf}(x)=\dfrac{1}{\sqrt{\pi}}\int_{-x}^{x}e^{-t^2}dt$ 叫误差概率积分, 它当 $x\to+\infty$ 有极限 1. 画出这个函数的图像, 并求出它的导数. 证明, 当 $x\to+\infty$ 时成立

$$\mathrm{erf}(x)=1-\frac{2}{\sqrt{\pi}}e^{-x^2}\left[\frac{1}{2x}-\frac{1}{2^2x^3}+\frac{1\cdot 3}{2^3x^5}-\frac{1\cdot 3\cdot 5}{2^4x^7}+o\left(\frac{1}{x^7}\right)\right].$$

怎么把这个渐近公式开拓成级数? 这个级数对某些 $x\in\mathbb{R}$ 收敛吗?

4. 道路的长依赖于用参数形式表示的运动规律吗?

5. 设您抓住了 1km 的橡皮绳的一端, 另一端固定在距您 1km 的地方. 从另一端有一个甲虫沿橡皮绳以 1cm/sec 的速度爬向您. 假定甲虫每爬过 1cm, 您就将橡皮绳拉长 1km. 试问, 甲虫能否爬到您手上? 如果能, 它大约需要多少时间?(奥库尼问题, 是萨哈罗夫提出来的.)

6. 计算质量在地球的重力场中移动的功, 并证明, 这个功只依赖于质量初始和最终位置的高度. 求逃逸地球重力场所需要的功和相应的 (第二) 宇宙速度.

7. 以单摆和双摆为例说明, 怎样在相应构形集合上引进局部坐标和邻域以及怎样产生出自然的拓扑, 从而把构形集合化成了力学系统的构形空间, 在所研究的情形能否将这个空间距离化?

8. $\mathbb{R}^n$ 中的单位球是紧集吗? 在 $C[a,b]$ 中如何呢?

9. 给定集合的一个子集叫做这个集合的 ε 网, 如果这个集的每一个点都有子集中的点与其距离小于 ε. 以 $N(\varepsilon)$ 表示给定集合的 ε 网中可能的最小点数. 试估计 $\mathbb{R}^n$ 中的线段, 正方形, 立方体和有界区域的 ε 熵 $\log_2 N(\varepsilon)$. 量 $\log_2 N(\varepsilon)/\log_2(1/\varepsilon)$ 当 $\varepsilon\to 0$ 时是否给出了所讨论的集合的维数的某种概念? 这样的维数可能等于 0.5 吗?

10. 设在 $\mathbb{R}^3$ 的单位球面 S 上, 温度 T 作为点的函数是连续的, 试问, 在球面上一定有温度的最小值点和最大值点吗? 当这两种点都存在时, 也存在它们的中间值点吗?

    当单位球 S 在空间 $C[a,b]$ 中取, 而温度在 $f\in S$ 的值是

$$T(f)=\left(\int_a^b |f|(x)dx\right)^{-1}$$

上面的结论还有哪些是对的?

11. a) 取 1.5 作为 $\sqrt{2}$ 的初始近似, 按牛顿方法进行两次迭代, 试问, 这两步计算各得到几位精确数字.

b) 作出求满足方程

$$f(x) = x + \int_0^x f(t)dt$$

的函数 f 的迭代程序.

多变量函数微分学

1. a) 在计算函数 $f(x,y,z)$ 在点 (x,y,z) 的值时, 如果坐标 (x,y,z) 的给出, 相应地, 有绝对误差 $\Delta x, \Delta y, \Delta z$, 试问, 函数值的相对误差 $\delta = |\Delta f| / |f|$ 是多少?

b) 如果房间的长 $x = 5 \pm 0.05\text{m}$, 宽 $y = 4 \pm 0.04\text{m}$, 高 $z = 3 \pm 0.03\text{m}$, 试问, 由此计算出的房间体积的相对误差是多少?

c) 线性函数值的相对误差能否与它的自变量值的相对误差相同?

d) 线性函数的微分能否与它本身一致?

e) 对于线性函数, 关系式 $f' = f$ 对吗?

2. a) 设有定义在圆盘上的二元函数, 它有一个偏导数在圆盘的每个点都等于零. 这是否意味着这个函数在圆盘上不依赖于相应的这个变量.

b) 如果把圆盘换成任意凸区域, 答案有变化吗?

c) 而若换成任意的区域呢?

d) 设 $x = x(t)$ 是一个点在平面 (或 $\mathbb{R}^n$) 内于时段 $t \in [a,b]$ 中的运动规律; $v(t)$ 是它的速度, 而 $C = \text{con v}\{v(t)|t \in [a,b]\}$ 是包含一切向量 $v(t)$ 的最小凸集 (也叫集合 $\{v(t)|t \in [a,b]\}$ 的凸包). 试证, 在 C 中存在向量 v 使 $x(b) - x(a) = (b-a)v$.

3. a) 设 $F(x,y,z) = 0$. 试问, 成立 $\dfrac{\partial z}{\partial y} \cdot \dfrac{\partial y}{\partial x} \cdot \dfrac{\partial x}{\partial z} = -1$ 吗? 试证, 它对依赖关系 $\dfrac{xy}{z} - 1 = 0$ (这相应于理想气体的克拉贝龙状态方程 $\dfrac{PV}{T} = R$) 是成立的.

b) 现设 $F(x,y) = 0$. 试问, $\dfrac{\partial y}{\partial x} \cdot \dfrac{\partial x}{\partial y} = 1$ 对吗?

c) 在一般依赖关系 $F(x_1, \cdots, x_n) = 0$ 的情形, 能作出什么结论?

d) 如果已知函数 $F(x,y)$ 在点 (x_0, y_0) 邻域中的泰勒公式的头几项, 这里 $F(x_0, y_0) = 0$, 而 $F'_y(x_0, y_0)$ 可逆, 试求由方程 $F(x,y) = 0$ 在 (x_0, y_0) 邻域中确定的隐函数 $y = f(x)$ 的泰勒展开式的头几项.

4. a) 验证: 椭球面 $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ 在其上一点 (x_0, y_0, z_0) 的切平面方程是

$$\frac{xx_0}{a^2} + \frac{yy_0}{b^2} + \frac{zz_0}{c^2} = 1.$$

b) 动点 $P(t) = \left(\dfrac{a}{\sqrt{3}}, \dfrac{b}{\sqrt{3}}, \dfrac{c}{\sqrt{3}}\right) \cdot t$ 于时刻 $t = 1$ 从椭球面 $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$ 出发. 设 $p(t)$ 是这个椭球面上在时刻 t 距 $P(t)$ 最近的点. 试求, 当 $t \to +\infty$ 时点 $p(t)$ 的极限位置.

5. a) 在具笛卡儿坐标 (x, y) 的平面上绘出函数 $f(x, y) = xy$ 的等值线以及曲线 $S = \{(x, y) | x^2 + y^2 = 1\}$. 利用所得之图, 对函数 f 在圆周 S 上的限制, 即函数 $f|_S$ 的极值问题进行全面的研究.

b) 寻找重力场中的质点在理想约束 (例如, $F_1(x, y, z) = 0, F_2(x, y, z) = 0$) 下的平衡位置, 可归结为求条件极值问题. 如果利用拉格朗日方法求解这个问题, 试问, 其中的拉格朗日乘数有何物理意义?

考试大纲

第一学期

分析引论 (数, 函数, 极限)

一元函数微分学

1. 实数,(上, 下) 有界数集, 完备性公理和集合上 (下) 界的存在性, 自然数集的无界性.
2. 与实数集 $\mathbb{R}$ 的完备性有关的基本引理 (闭区间套定理, 有限覆盖定理, 极限点定理)
3. 序列的极限及其存在性的柯西准则, 单调序列极限存在性判别法.
4. 级数及其和, 几何级数, 级数收敛性的柯西准则和必要条件, 调和级数, 绝对收敛.
5. 非负项级数收敛性判别法, 比较定理, 级数 $\zeta(s)=\sum\limits_{n=1}^{\infty} n^{-s}$.
6. 对数的思想和数 e, 函数 $\exp(x)$ 及其幂级数表达式.
7. 函数的极限, 基本的极限过渡的基底, 在任意基底下函数极限的定义及其在具体情况下的具体形式, 无穷小函数和它们的性质, 函数终极性质的比较, 渐近公式以及符号 $o(\cdot), O(\cdot)$ 的基本运算.
8. 极限过渡与代数运算, $\mathbb{R}$ 中的序关系之间的相互联系. 当 $x\to 0$ 时 $\dfrac{\sin x}{x}$ 的极限.
9. 复合函数和单调函数的极限, 当 $x\to\infty$ 时 $(1+\dfrac{1}{x})^x$ 的极限.
10. 函数极限存在性的柯西准则.
11. 函数在一点的连续性, 连续函数的局部性质 (局部有界性, 保号性, 算术运算, 复合函数的连续性). 多项式, 有理函数和三角函数的连续性.
12. 连续函数的整体性质 (中间值定理, 最大值, 等度连续性).
13. 单调函数的间断点, 反函数定理, 反三角函数的连续性.

14. 运动律, 小时间间隔中的位移, 瞬时速度向量, 轨迹及其切线, 函数在一点可微的定义, 微分及其定义域和值域, 微分的唯一性, 实变量的实值函数的导数及其几何意义, 函数 $\sin x, \cos x, e^x$, $\ln|x|, x^\alpha$ 的可微性.
15. 可微性和算术运算, 多项式, 有理函数, 正切和余切函数的微分.
16. 复合函数和反函数的微分, 反三角函数的导数.
17. 函数的局部极值, 可微函数内部极值点的必要条件 (费马引理).
18. 罗尔定理, 拉格朗日和柯西有限增量 (中值) 定理.
19. 具柯西和拉格朗日形式的余项的泰勒公式.
20. 泰勒级数, 函数 $e^x, \cos x, \sin x, \ln(1+x), (1+x)^\alpha$(牛顿二项式) 的泰勒展开式.
21. 局部泰勒公式 (带有佩亚诺余项).
22. 可微函数的单调性与其导数的非负性的关系, 局部极值存在性与不存在性的用一阶、二阶和高阶导数表述的充分条件.
23. 凸函数, 凸性的微分条件, 凸函数的图像与其切线的相对位置.
24. 凸函数的一般詹生不等式, 对数的凸性, 柯西, 杨格, 赫尔德和闵可夫斯基的经典不等式.
25. 复数的代数记法和三角记法, 复数序列的收敛性和复数项级数的收敛性, 柯西准则. 复数项级数的绝对收敛性及其充分性判别法, 极限 $\lim\limits_{n\to\infty}(1+\frac{z}{n})^n$.
26. 幂级数的收敛圆和收敛半径, 函数 $e^z, \cos z, \sin z (z\in\mathbb{C})$, 欧拉公式初等函数的相互联系.
27. 作为现象的数学模型的微分方程, 例子, 待定系数法和欧拉折线法.
28. 原函数, 求原函数的基本方法 (函数和的逐项积分, 分部积分, 变量替换). 基本初等函数的原函数.

第二学期

一元函数的积分

多变量函数的微分学

1. 区间上的黎曼积分, 可积性的必要条件, 零测度集, 它们的一般性质, 例子, 函数黎曼可积的勒贝格判别法 (叙述). 可积函数空间和可积函数的运算.
2. 积分的线性性质, 可加性和一般估计.
3. 实值函数积分的估计,(第一) 中值定理.
4. 变上限积分性质, 连续函数的原函数的存在性, 广义原函数及其一般形式.
5. 牛顿–莱布尼茨公式, 积分中的变量替换.
6. 定积分中的分部积分法, 具积分余项的泰勒公式, 第二中值定理.
7. 有向区间的可加函数和积分, 积分应用的一般模式, 例子: 道路的长 (及其关于参数的不变性), 曲边梯形的面积, 旋转体的体积, 功, 能.
8. 黎曼–斯蒂尔切斯积分. 化成黎曼积分的条件, 狄拉克 δ 函数的奇异性, 广义函数的概念.

9. 反常积分的概念, 典则积分, 研究反常积分收敛性的柯西准则和比较定理, 级数收敛性的积分判别法.

10. 局部线性化, 例子: 瞬时速度和位移; 小振动摆运动方程的简化; 在自变之微小变化条件下, 计算 $\exp(A), A^{-1}, \det(E), < a, b >$ 的线性修正 (这里, A 是可逆方阵, E 是单位矩阵, a 和 b 是向量, $< \cdot, \cdot >$ 是内积).

11. 向量的范数 (长, 模), 向量空间的重要例子, 线性连续算子空间 $L(X, Y)$ 及其中的范数, 线性算子的连续性及其范数的有限性.

12. 函数在一点的可微性, 微分, 其定义域和值域, 映射 $f: \mathbb{R}^m \to \mathbb{R}^n$ 的微分的坐标表示, 可微性、连续性和偏导数的存在性之间的关系.

13. 复合函数和反函数的微分, 适用于映射 $f: \mathbb{R}^m \to \mathbb{R}^n$ 的各种情形的微分法则的坐标表示.

14. 方向导数和梯度. 几何和物理中应用梯度的例子 (函数的等值线, 坡度, 切平面; 势场; 理想流体动力学的欧拉方程, 伯努利定律, 机翼的作法).

15. 齐次函数及欧拉关系, 量纲分析方法.

16. 有限增量定理及其几何、物理意义, 应用的例子 (用偏导数表述的可微性充分条件; 在区域中函数取常值的条件).

17. 高阶导数及其对称性.

18. 泰勒公式.

19. 函数的极值 (内部极值的必要条件和充分条件).

20. 压缩映射, 毕卡–巴纳赫不动点原理.

21. 隐函数定理.

22. 反函数定理, 曲线坐标和拉直变换, $\mathbb{R}^n$ 中的 k 维光滑曲面及其切平面, 定义曲面的方法及相应的切空间方程.

23. 秩定理和函数相关性.

24. 条件极值 (必要性判别法). 拉格朗日方法的几何的、代数的和物理的解释.

25. 条件极值的充分性判别法.

26. 度量空间, 例子. 开子集和闭子集, 点的邻域. 诱导度量, 子空间, 拓扑空间, 点的邻域, 分离性 (豪斯多夫公理). 在子集上的诱导拓扑. 集合的闭包和相对闭子集的概念.

27. 紧集及其绝对性, 紧集是闭集, 紧集的闭子集是紧集, 紧集套. 度量空间中的紧集, ε 网. 度量空间中紧性的判别法及其在 $\mathbb{R}^n$ 中的具体化.

28. 完备度量空间. $\mathbb{R}, \mathbb{C}, \mathbb{R}^n, \mathbb{C}^n$ 以及连续函数空间 $C[a, b]$ 关于一致收敛的完备性.

29. 拓扑空间中映射的连续性的判别. 连续映射保持集合的紧性和连通性. 关于连续函数的有界性, 最大值和中间值的经典定理. 在度量空间的紧集上定义的连续函数必一致连续.

参考文献

I. 经典文献

1. 原著

Ньютон И.

a. Математические начала натуральной философии. Пер. с лат. В кн.: Крылов А. Н. Собрание трудов. Т.7.—Л.–М.: Изд-во АН СССР,1936, с. 57–662.

b. Математические работы.—М.–Л.: ОНТИ, 1937.

Лейбниц Г. В. Избранные отрывки из математических сочинений. *Успехи матем. наук*, 1948. 3(1), 165–205.

2. 重要专著

Эйлер Л.

a. Введение в анализ бесконечных. В 2-х т.—М.: Физматгиз, 1961.

b. Дифференциальное исчисление.—М. Л.: Гостехиздат, 1949.

c. Интегральное исчисление. В 3-х т.—М.: Гостехиздат, 1956–1958.

Коши О. Л.

a. Алгебраический анализ. Лейпциг: Бэр и Хэрманн, 1864.

b. Краткое изложение уроков о дифференциальном и интегральном исчислении. — СПб.: Имп. Акад. наук, 1831.

II. 教科书[①]

Архипов Г. И., Садовничий В. А., Чубариков В. Н. Лекции по математическому анализу. —М.:Высшая школа, 2000.

Ильин В. А., Садовничий В. А., Сендов Б. Х. Математический анализ. В2-х ч. Изд. 2-е, перераб. М.: Изд-во Моск. ун-та, 1985, 1987.

Камынин Л. И. Курс математического анализа. В 2-х ч. —М.: Изд-во Моск. ун-та, 1993, 1995.

Кудрявцев Л. Д. Курс математического анализа. В 3-х т. —М.: Высшая школа, 1988, 1989.

Никольский С. М. Курс математического анализа. В 2-х т. — М.: Наука, 1990.[②]

III. 教学参考书

Виноградова И. А., Олехник С. Н., Садовничий В. А. Задачи и упражнения по математическому анализу. —М.: Изд-во Моск.ун-та, 1988.

Демидович Б. П. Сборник задач и упражнений по математическому анализу. —М.: Наука, 1990.[③]

Макаров Б. М., Голузина М. Г., Лодкин А. А., Подкорытов А. Н. Избранные задачи по вещественному анализу. —М.: Наука, 1992.

Решетняк Ю. Г. Курс математического анализа. —Новосибирск:Издво Инс-та матем. Ч. I, книги 1 и 2, 1999. Ч. II, книги 1 и 2, 2000, 2001.

Рудин У. Основы математического анализа. Изд.2-е. —М.: Мир, 1976.[④]

Шилов Г. Е.

a. Математический анализ. Функции одного переменного. Ч. 1 2. —М.: Наука, 1969.

b. Математический анализ. Функции нескольких вещественных переменных. В 3-х ч. —М.: Наука, 1972.

---

①在这一部分援引的书，是苏联高教部核 二组，俄罗斯科学部高等学校委员会或俄罗斯联邦教育部向数学、应用数学、力学、应用数学与信息科学等专业的大学生推荐的教材.

②中译本：尼柯尔斯基著. 数学分析教程. 第一卷第一分册，第二分册. 刘远图，郭思旭，高尚华译. 北京：高等教育出版社，1980，1981.

③中译本：Б. П 吉米多维奇著. 数学分析习题集 第一版. 李荣涷译. 北京. 高等教育出版社，1958.

④中译本· W. 卢丁著. 数学分析原理. 上、下册. 赵慈庚，蒋铎译. 北京：人民教育出版社，1979.

Фихтенгольц Г. М. Курс дифференциального и интегрального исчисления. В 3-х т. Изд. 7-е, стереотип. —М.: Наука, 1969.[①]

IV. 补充文献

Александров П. С., Колмогоров А. Н. Введение в теорию функций действительного переменного. —М.: ГТТИ, 1938.

Альберт Эйнштейн и теория гравитации: Сб. статей. К 100-летию со дня рождения. —М.: Мир, 1979.

Арнольд В. И.

a. Гюйгенс и Барроу,Ньютон и Гук—первые шаги математического анализа и теории катастроф, от эвольвент до квазикристаллов. —М.: Наука, 1989.

b. Математические методы классической механики. —М.:Наука, 1989.

Бурбаки Н. Очерки по истории математики. —М.: Изд-во иностранной литературы, 1963. (В частности, статья《Архитектура математики》.)

Валле-Пуссен Ш. Ж. Курс анализа бесконечно малых. В 2-х т.Л. —М.: ГТТИ, 1933.

Вейль Г. Математическое мышление. —М.: Наука, 1989.

Гелбаум Б., Олмстед Дж. Контрпримеры в анализе. —М.: Мир, 1967.[②]

Гельфанд И. М. Лекции по линейной алгебре. —М.: Добросвет, МЦНМО, 1998.

Дубровин Б. А., Новиков С. П., Фоменко А. Т. Современная геометрия: Методы и приложения. —М.: Наука, 1986.

Дьедонне Ж. Основы современного анализа. —М.: Мир, 1964.

Зельдович Я. Б., Мышкис А. Д. Элементы прикладной математики. —М.: Наука, 1967.

Зорич В. А. Анализ. [Записки лекций для студентов Математического колледжа НМУ и механико-математического ф-та МГУ.] В 3-х вып. Вып. I. Лекции. 5–7: Дифференциал. Вып. II. Лекция 8: Теорема о неявной функции. Вып. III. Лекции 9–11: Приложения теоремы о неявной функ-

①中译本· 微积分学教程 (第 8 版). Г. М. 菲赫金哥尔茨著. 北京: 高等教育出版社, 2006.
第一卷 杨弢亮, 叶彦谦译, 郭思旭校.
第二卷. 徐献瑜, 冷生明. 梁文骐译. 郭思旭校.
第三卷. 路见可, 余家荣. 吴亲仁译, 郭思旭校.

②中译本: [美]B. R. 盖尔鲍姆, J. M. H. 奥姆斯特德著. 分析中的反例. 高枝译. 上海: 上海科学技术出版社, 1980.

ции. М.: Изд-во механико-математич. ф-та МГУ, 1995.

Картан А. Лифференциальное исчисление. Дифференциальные формы. М.: Мир, 1971.

Колмогоров А. Н., Фомин С. В. Элементы теории функций и функционального анализа. Изд. 4-е, перераб. —М.: Наука, 1976.①

Кострикин А. И., Манин Ю. И. Линейная алгебра и геометрия. —М.: Наука, 1986.

Кириллов А. А. Что такое число? —М.: Наука, 1993.

Курант Р. Курс дифференциального и интегрального исчисления. В 2-х т. —М.: Наука, 1970.②

Ландау Э. Основы анализа. —М.: Изд-во иностранной литературы, 1947.③

Манин Ю. И. Математика и физика. —М.: Знание, 1979. —(Новое в жизни, науке, технике. Серия: Математика, кибернетика; № 12.)

Милнор Дж. Теория Морса. —М.: Мир, 1965. —(Библиотека сборника 《Математика》.)

Нарасимхан Р. Анализ на действительных и комплексных многообразиях. —М.: Мир, 1971.

Полиа Г., Сеге Г. Задачи и теоремы из анализа. В 2-х ч. Изд. 3-е. —М.: Наука, 1978.④

Понтрягин Л. С. Обыкновенные дифференциальные уравнения. —М.: Наука, 1974.

Пуанкаре А. О науке. —М.: Наука, 1990.

Спивак М. Математический анализ на многообразиях. —М.: Мир, 1971.⑤

Уиттекер Э. Т., Ватсон Дж. Н. Курс современного анализа. В 2-х ч. Изд. 2-е. —М.: Физматгиз, 1962–1963.

Успенский В. А. Что такое нестандартный анализ? —М.: Наука, 1987.

Фейнман Р., Лейтон Р., Сэндс М. Фейнмановские лекции по физике. Т. I: Современная наука о природе. Законы механики. —М.: Мир, 1965.

Халмош П. Конечномерные векторные пространства. —М.: Наука, 1963.

---

①中译本 A H. 柯尔莫戈洛夫, C B. 佛明著 函数论与泛函分析初步 (上、下册). 段虞荣, 郭思旭, 郑洪深译 北京: 高等教育出版社, 1992.

②中译本: R. 柯朗, F. 约翰著. 微积分和数学分析引论.
第一卷. 张鸿林, 周民强译. 北京: 科学出版社, 1979.
第二卷第一分册. 林建祥, 刘婉如等译. 北京: 科学出版社, 1985.

③中译本 艾 · 兰道著. 分析基础. 刘绂堂译. 北京: 高等教育出版社, 1958.

④中译本: G. 波利亚, G. 舍贵著. 数学分析中的问题和定理. 第一卷, 第二卷 张奠宙, 宋国栋等译. 上海: 上海科学技术出版社, 1985.

⑤中译本: [美] 斯皮瓦克著. 流形上的微积分. 齐民友, 路见可译. 北京: 科学出版社, 1980.

Шварц Л. Анализ. В 2-х т. —М.: Мир, 1972.

Эйнштейн А. Собрание научных трудов. Том IV. —М.: Наука,1967.(В том числе статьи 《Мотивы научного исследования》(с. 39–41)и《Физика и реальность》(с. 200–227).)

名词索引

δ 邻域, 60, 367, 368
N_0 的标准模型, 24
R^m 中的区域, 378
m 维曲面, 420
n 次迭代, 150
n 阶相切, 162
n 质点系的构形空间, 12
$y = f(x)$ 叫做函数 f 在元素 x 上的值, 18
$(m-1)$ 维曲面, 432
(按冯 · 诺依曼 (Neumann) 的建议) 建立自然数 N_0 的标准模型, 24
(乘法) 群, 30

A

阿基米德原理, 56
鞍形临界点, 419

B

摆的振动, 347
摆的振动周期, 403
半径
  临界半径, 265
  收敛半径, 243
半衰期, 264
包含关系, 6
闭集, 369
闭球, 367–369
闭区间套, 59
闭区间套引理, 60
必要条件, 2
变换, 10
  伽利略变换, 187
  阿贝尔变换, 319
  对合变换, 236
  勒让德变换, 236, 438, 440
  洛伦兹变换, 21, 186, 187
  伽利略变换, 21, 184, 185
变上限的积分, 324
表
  简单的不定积分表, 279
并集, 7
波尔察诺–柯西中间值定理, 141
伯努利不等式, 76
补集, 7
不变集, 21
不等式
  阿达马不等式, 480
  伯努利不等式, 54, 216
  赫尔德不等式, 217, 224, 323
  积分的赫尔德不等式, 323

柯西–布尼雅可夫斯基不等式, 323
二角形不等式, 46, 366, 367, 383
施瓦茨不等式, 323
杨格不等式, 216, 236, 351
詹生不等式, 223, 323
闵可夫斯基不等式, 217, 218, 323, 366, 383
不动点, 150
部分
复数 z 的实部和虚部, 239
积分的有理部分, 293
小数部分, 44
整数部分, 44

C

参数, 342
分划 P 的参数, 299
自然参数, 347
自然参数化, 347
测地线, 12
差集, 7
场
函数的梯度场, 399
力场, 400
势场, 399
位场, 399
向量场, 397, 399
常数
欧拉常数, 131
普朗克常数, 48
时间常数, 274
万用气体常数, 263
万有引力常数, 48
长
道路的长度, 338
光滑道路的长度的定义, 339
曲线的长, 12
椭圆的周长, 343
乘法, 31
承载子, 338, 339, 342, 380
充分条件, 2
传递性, 17, 18, 399
存在量词, 6

D

代数封闭性, 253
待定系数法, 258
单调性, 219, 314
单位
虚单位, 238
弹性系数, 345
导数, 159–161
对数导数, 177
复变函数 $f(z)$ 在点 z_0 处的导数, 251
高阶导数, 187
高阶偏导数, 406, 407
偏导数, 387
沿给定方向的导数, 397
沿向量 v 的导数, 396
右导数, 236
左导数, 236
道路, 338, 376, 378, 399, 420, 477
光滑道路, 458
简单闭道路, 338, 339
简单道路, 339
等价
f 与 g 等价, 125
等价关系, 19
笛卡儿积, 8
狄利克雷函数, 312
点
边界点, 368
不动点, 21, 150
插值节点, 211
第二类间断点, 139
第一类间断点, 138, 146
非退化临界点, 454, 457
拐点, 222, 438
极限点, 60, 368

间断点, 137
局部极大(或极小)值点, 191
局部极值点, 191, 411
可去间断点, 138
临界点, 411, 454, 457, 477, 479, 480
内点, 368
内极值点, 192
切比雪夫交错点, 152
外点, 368
稳定点, 412, 480
严格局部极大(小)值点, 191
点列, 372
典则变量, 440
调和平均值, 92
迭代法, 15
n 次迭代, 26
定理
阿贝尔第一定理, 243
比较定理, 84, 85
达布定理, 209, 312
戴德金定理, 54, 56
代数基本定理, 255
德拉瓦莱–布森定理, 152
第二中值定理, 360
第一中值定理, 317, 333
反函数定理, 442
关于反函数的定理, 145
关于基数的定理(康托尔), 23
关于有限增量的拉格朗日定理, 193, 194
积分的第二中值定理, 319, 322
拉格朗日定理, 204
拉格朗日中间值定理, 194
量纲理论的 Π–定理, 403
罗尔定理, 193, 423
切比雪夫定理, 152
施略德–伯恩斯坦定理, 22, 26
完备公理, 54
魏尔斯特拉斯定理, 74
隐函数定理, 426, 435
隐函数定理的最简单情形, 428
有限增量定理, 194, 404
折射定律, 480
中值定理, 403, 404
定律
光的折射定律, 215
开普勒定律, 154, 403
克拉贝龙定律, 263, 441
牛顿第二定律, 266
牛顿定律, 154, 270, 399
欧姆定律, 21
斯涅耳定律, 215
速度加法定律, 184–186
万有引力定律, 154, 266, 403
度量
$\mathbb{R}^m$ 中的距离, 365
$x, y \in \mathbb{R}$ 之间的距离, 46
度量空间, 366
集合 E_1 与 E_2 之间的距离, 371
完备度量空间, 372
度量, 距离
f 和 g 之间的距离, 151
对称性, 17, 399
对基 $\mathcal{B}$ 为无穷小, 114
对基 $\mathcal{B}$ 最终为常数, 114
对基 $\mathcal{B}$ 最终有界, 114
对数, 108–110, 170
对数刻度, 177
积分对数, 283, 362
对数函数, 108, 109, 111
对于基 $\mathcal{B}$, 函数 f 渐近于函数 g, 125
多项式
拉格朗日插值多项式, 210, 334
勒让德多项式, 334
切比雪夫多项式, 152
泰勒多项式, 202, 205
最佳逼近多项式, 151
多重指标, 422

E

二次平均值, 92
二维曲面, 417, 418, 432
二项式
牛顿二项式, 54, 188, 200, 201
微分二项式, 294

F

发散数列, 67
法则
洛必达法则, 225
反对称, 19
反对称性, 18
反函数, 148
反身性, 17, 18, 399
反双曲函数, 181
范畴的, 32
范数, 383, 385
泛函, 10, 12, 314
方程
变量分离的方程, 296
简谐振动方程, 271
拉普拉斯方程, 421
欧拉方程, 400
欧拉折线法, 269
微分方程, 158, 260, 263, 269, 296
热传导方程, 421
方程组
哈密顿方程组, 440
欧拉–拉格朗日方程组, 440
方法
奥斯特洛格拉德斯基方法, 292
待定系数法, 269
拉格朗日乘数法, 469
量纲方法, 402
穷尽法, 299
梯度方法, 397
最小二乘法, 423
放射衰变, 264, 274
费马引理, 191
分部积分法, 280
分部积分公式, 327
分段光滑道路, 339
分划, 299
带标志点的分划, 300
分划 P 的区间, 299
分解
把真分式表示成所谓简分式的和, 257
分解微分同胚为最简形式的复合, 451
函数 f 在点 x_0 处的泰勒级数, 201
函数 $f(x)$ 在点 x_0 处的 n 阶泰勒多项式, 196
辐角, 239
辐角的分支 (或主支), 239
复合
关系 $\mathcal{R}_1, \mathcal{R}_2$ 的复合 $\mathcal{R}_2 \circ \mathcal{R}_1$, 19
复合函数的微分法, 175
复合映射, 373
复平面, 239
复数辐角, 239, 240

G

高斯积分, 361
割线, 162, 164
根, 254, 256
复数 a 的 n 个不同的根, 240
实根, 151
数 1 的所有的 n 次根, 259
重根, 256, 258
公理, 23
并公理, 23, 27
策梅洛公理, 24
乘法公理, 30, 33
对公理, 23, 27
分出公理, 23, 27
公理系统, 23
加法公理, 29, 32
容积公理, 23, 27
完备(连续)公理, 45

完备公理, 31, 36, 37, 43, 45, 55–59, 61
完备性公理, *116*
无穷公理, 24, 27
序公理, 31, 34
选择公理, 24
置换公理, 24
子集之集的公理, 24
公理系统, 45, 61
策梅洛–弗兰克尔 (Zermelo-Fraenkel) 公理系统, 24
公理系统的范畴性, 32
公理系统的无矛盾性, 32
集合论公理系统, 24
实数集的公理系统, 29, 43, 57
公式
埃尔米特插值公式, 211
奥斯特洛格拉德斯基公式, 292, 293
波内公式, 322
带积分余项的泰勒公式, 327
分部积分公式, 281
复数的代数形式和三角形式, 239
局部泰勒公式, 203
矩形公式, 335, 336
具有拉格朗日形式的泰勒公式余项, 410
柯西–阿达马公式, 243
拉格朗日插值公式, 211
拉格朗日余项的多重指标记号的泰勒公式, 422
莱布尼茨公式, 188, 189
麦克劳林公式, 198
牛顿–莱布尼茨公式, 299, 330
欧拉公式, 245
抛物形公式, 335, 336
佩亚诺形式余项的泰勒公式, 410
气压公式, 262
求积公式, 336
泰勒公式, 196, 409, 412, 422
泰勒公式余项, 336
梯形公式, 335, 336
韦达公式, 132
辛普森公式, 335, 336
余项的拉格朗日公式, 198
棣莫弗公式, 240, 248
关系, 4, 17
包含关系, 18
出发域, 17
到达域, 17
等价关系, 17, 19, 21, 399
等势关系, 21
关系 $\mathcal{R}$ 的定义域, 17
关系 $\mathcal{R}$ 的值域, 17
函数的图像, 18
函数关系, 18
序关系, 18
转置关系, 19
关系的定义域, 17
关于 X 中基 $\mathcal{B}$ 的极限, 371
关于数学命题的逻辑结构, 24
光滑道路, 339

H

哈密顿 (Hamilton) 方程组, 440
函数, 10, 16, 18, 19
c_0–等高集, 471
c_1–等高集, 471
f 是当 $E \ni x \to a$ 时的无穷小, 99
n 次齐次函数, 401
凹函数, 219
常值函数, 97
处处不可导的连续周期函数, 171
单调函数, 120, 121
等高集, 433, 471
狄利克雷函数, 139, 312
调和函数, 421
定向区间的可加函数, 316, 336, 337
对数函数, 108, 112
多变量函数, 365
多变量可微函数, 386
多变量连续函数, 376

反函数, 145, 146
符号函数, 95
复变量 z 的复值函数, 248
函数 f 的等高集(c–等高面), 399
函数的复合, 14
函数的图像, 19
函数等高集, 467
函数在点 $z_0 \in \mathbb{C}$ 处解析, 252
函数作为子集 $f \subset X \times Y$, 19
解析函数, 201, 260
局部 n 次齐次的, 401
可积函数, 301
可微复变函数, 251
可微函数, 154, 159
拉格朗日函数, 469, 470, 474–476
黎曼函数, 139, 312
力函数, 399
连续复变函数, 250
连续函数, 133
幂函数, 111, 112
面积余切, 182
面积正切, 182
三角函数, 246, 340
上凸函数, 219
上有界函数, 98
势函数, 399, 400
双曲函数, 246
双曲余切, 182
双曲余弦, 180
双曲正切, 182
双曲正弦, 180
凸函数, 219
无穷可微函数, 251
下凸函数, 219
下有界函数, 98
严格凸函数, 218, 219
一致连续函数, 143, 144
以 $a(0 < a, a \neq 1)$ 为底的对数函数, 108
隐函数, 183, 189, 434, 435
隐函数定理, 426
有界函数, 98
圆函数, 246
在集合 E 上的连续的实函数, 136
在集合 E 的极限点 $x \in E$ 处是可微的函数, 159
指数函数, 103, 108, 112, 246–248, 267
周期函数, 171, 271, 331, 334
自然对数, 108
最终上有界, 98
最终为常数函数, 122
最终为常值函数, 97
最终下有界函数, 98
最终有界, 98
最终有界函数, 122
作为关系的函数、函数的图像, 16
函数 f 在点 a 的微分, 158
函数 f 在集合 A 上的缩小或限制, 10
函数 $f: E \to \mathbb{R}$ 在点 $a \in E$ 的振幅, 135
函数 $f: E \to \mathbb{R}$ 在集合 E 上连续, 136
函数(映射), 10
函数 $f: E \to \mathbb{R}$ 在点 $a \in E$ 连续, 134
函数 $f: E \to \mathbb{R}^n$ 在点 $a \in E$ 的振幅, 377
函数 $f: X \to \mathbb{R}^n$ 在 E 上的振幅, 373
函数的变元或自变量, 10
函数的出发域, 10
函数的到达域, 10
函数的定义域, 10
函数的渐近行为, 128, 203
函数的幂级数表示, 解析性, 248
函数的因变量, 10
函数的值集, 10
函数的值域, 10
函数独立, 450, 457
函数相关, 450, 457
函数相关的, 450
函数芽, 152
函数在集 X 上的振幅, 135
函数值, 10

赫尔德不等式, 217
黑塞矩阵, 439
环
　　连续函数环, 153
　　连续函数芽的环, 153

J

基, 113, 114
　　分划集的基, 300
　　分析中最常用的基, 112
　　集合 X 中的基, 112
　　滤子基, 113
基本点列, 372
基本列, 372
基底, 381, 385
基数, 22
积
　　集的直积, 23
　　级数的积, 245
　　数量积, 384, 385
　　无穷乘积, 131
积分, 299, 301, 324
　　不定积分, 276, 277, 280
　　超椭圆积分, 291
　　第二类椭圆积分, 292
　　第三类椭圆积分, 292
　　第一类全椭圆积分, 363
　　第一类椭圆积分, 292, 363
　　定积分, 326, 327
　　发散的反常积分, 352
　　反常积分, 351–354
　　菲涅耳积分, 296, 332, 363
　　概率误差积分, 363
　　积分对数, 296
　　积分双曲余弦, 296
　　积分双曲正弦, 295
　　积分余弦, 283, 295, 362
　　积分正弦, 283, 295, 362
　　积分指数, 295, 363
　　绝对收敛的反常积分, 356
　　黎曼积分, 301, 351
　　欧拉 泊松积分, 296
　　上达布积分, 313
　　收敛的反常积分, 352
　　条件收敛的反常积分, 359
　　椭圆积分, 291, 292, 295
　　下达布积分, 313
　　向量值函数的积分, 313
积分的第一中值定理, 318
积分的下限和上限, 301
积分的线性性, 314
积分法
　　不定积分法, 277, 278
积分和, 299–301, 304
　　上积分和, 306
　　下达布和, 306
积分上限, 316
积分下限, 315, 316
积分余项, 410
极大(小)值
　　局部极大值(局部极小值), 411
　　用高阶导数表达的极值的充分条件, 215
　　用一阶导数表达的极值的充分条件, 214
极大(小)值条件极值存在或不存在的充分条件, 472
极大元, 56
极大值, 424
极限
　　部分极限, 80
　　复合函数极限, 373
　　关于基 B 的极限, 113
　　关于基 B_X 的极限, 117
　　关于基的极限, 113, 114
　　函数的极限, 93
　　函数极限, 97, 112
　　数列 $\{x_k\}$ 的上极限, 78
　　数列 $\{x_k\}$ 的下极限, 78
　　数列 $\{x_n\}$ 的极限, 67
　　数列极限, 68, 72

极小曲面, 440
极小元, 56
极小值, 423
集合, 4, 5, 23
  闭集, 367–370
  不变集, 21
  超越数集, 43
  代数数集, 43, 62
  归纳集, 24, 38, 55
  后续集, 24
  开集, 367, 368
  康托尔集, 64, 65, 313
  可数集, 61, 62
  空集, 23
  空子集, 23
  连通集, 378–380
  零测度集, 311, 313
  路连通集, 378
  偏序集, 31
  上(下)有界集, 36
  实数集, 29
  稳定集, 21
  无界集, 46, 373
  无穷集, 22, 23, 55, 61, 62
  线性序集, 18, 31
  有界集, 36, 370
  有理数集, 41
  有穷集, 22
  整数集, 40
  至多可数的, 62
  至多可数集, 62
  子集, 23
  自然数集, 38
集合 E 的特征函数, 12
集合 $E \subset \mathbb{R}^m$ 的直径, 370
集合的运算, 9
集合论公理系统, 5
集合运算, 7
级数, 81
  调和级数, 83
  非负项级数, 84
  复数级数绝对收敛的定义, 242
  复数级数收敛柯西准则, 242
  复数项级数, 241
  级数的和, 82
  级数发散, 82
  级数收敛, 82
  绝对收敛级数, 83
  幂级数, 242
级数的部分和, 82
级数的项重排, 83
级数或无穷级数, 81
几乎处处, 311, 313, 323
几乎所有点, 311
记数法的基, 51
伽利略变换, 11
夹角
  曲线在这点的夹角, 478
  曲线之间的夹角, 478
  曲线之间的交角, 478
  向量 x 与 y 之间的夹角, 385
加法, 31
加速度, 154–156, 158, 171, 189
加速度向量, 237
渐近线, 231, 234
  倾斜渐近线, 229, 232
  竖直渐近线, 229
  水平渐近线, 229
渐近行为, 121, 122
  对基 B 的渐近行为, 122
  函数的渐近行为, 121, 123
交集, 7
结构
  $\mathbb{R}^m$ 的欧几里得结构, 384
界
  上界, 37
  有界数列, 77
紧集, 143, 369, 370, 379, 380
  $\mathbb{R}^m$ 中的紧集, 369–371

局部极大(或极小)值, 191
局部极值, 191, 411
局部齐次函数, 401
局部抻直, 445
矩阵
黑塞矩阵, 453, 479
雅可比矩阵, 389, 392, 411, 446, 448
具有拉格朗日余项的泰勒公式, 204, 211
具有佩亚诺余项的泰勒公式, 204
距离, 366
绝对值, 46
均方值, 92

K

开集, 369
开球, 367–369
康托尔海涅关于一致连续的定理, 144
康托尔集, 313
康托尔零测度集, 313
柯西–黎曼方程组, 457
柯西数列, 72
可加性, 314
可数集, 61, 62
可微函数, 164, 172
空集, 9
空间
n 质点系的相空间, 13
度量空间, 366
空间 $\mathbb{R}^m$, 365
欧几里得空间, 385
切空间, 386, 462–465, 478
完备度量空间, 372
完备空间, 116
向量空间 $\mathcal{R}[a,b]$, 308
空子集, 6
扩张或延拓, 10

L

莱布尼茨公式, 189
勒贝格判别法, 313
勒让德变换, 236, 438–440
黎曼函数, 312
黎曼可积函数, 301
理想, 153
极大理想, 153
连续函数芽的环的理想, 153
力场, 399
连分数, 91
无穷连分数, 91, 92
无穷连分数的近似分数, 91
有限连分数, 91
有限链式分数, 91
连续模, 150
连续统, 63
连续性, 133, 143, 371, 376
列, 67
临界点, 412
临界点的指标, 457
临界质量, 265
邻域, 46, 60, 94, 368, 373, 376, 377
δ 邻域, 46
去心邻域, 94
零测度的康托尔集, 313
零测度集, 310, 311
刘维尔 (Liouville) 定理, 55
流线, 400
滤子极限, 113
逻辑符号, 1
逻辑运算, 3, 9
落体, 266
洛伦兹变换, 11

M

米歇尔斯基, 262
幂级数, 243
面积
曲边梯形的面积, 343, 344
椭圆的面积, 344
面积正弦, 180
模, 46

复数 z 的模, 239
连续模, 151

N
能
动能, 13
势能, 13
总能量, 13
逆映射的微分法, 397
牛顿–莱布尼茨, 329
牛顿定律, 191

O
欧几里得辗转除法, 55, 91
欧拉–泊松积分, 361
欧拉–拉格朗日方程组, 440
欧拉替换, 290

P
判别法
常值函数判别法, 195
达布判别法, 313
达朗贝尔比较检验法, 200
达朗贝尔检验法, 87
单调数列的极限存在准则, 74
杜布瓦雷蒙判别法, 313
反常积分收敛性的柯西判别法, 356
非负级数的判敛准则, 84
复数级数收敛的柯西准则, 241
函数单调性检验法, 195
函数极限存在的柯西准则, 115, 373
函数黎曼可积性的勒贝格判别法, 310, 313
积分收敛性的阿贝尔–狄利克雷准则, 360
级数绝对收敛性的高斯检验法, 132
级数收敛的柯西准则, 82
柯西根值检验法, 86, 87
数列收敛的柯西准则, 72, 372
魏尔斯特拉斯比较检验法, 85, 86
西尔维斯特准则, 414
判断法
在 E 上不降的, 120
在 E 上不增的, 120
在 E 上递降的, 121
在 E 上递增的, 120
偏序关系, 18
平均
次平均值, 92
调和平均, 235
二阶平均, 235
积分平均值, 333
几何平均, 224
加权的 t 阶平均, 235
平均函数, 334
算术平均, 224, 235

Q
齐次函数, 401
齐次函数的欧拉恒等式, 401
切空间, 464
切平面, 418, 419, 426, 462–465
切线, 162, 426
球, 367
球面, 368
球面投影, 478
区间, 45
m 维区间, 369
半开区间, 45
闭区间, 45
开区间, 45
无界区间, 46
区间 ab 的长, 46
曲边梯形, 343, 344
曲率, 237
绝对曲率, 237
曲率半径, 237
曲面, 417, 419, 420, 432, 433, 458–463, 465–468, 471–473, 477
曲线, 12, 339
参数化曲线, 339

简单闭曲线, 339
曲线的参数形式, 342
去心 δ 邻域, 94
去心邻域, 373
全称量词, 6
确界
上确界, 37
群, 30
阿贝尔群, 30, 40
加群, 30
交换群, 30

S

三角形不等式, 218
上界, 36, 37, 43, 72
上确界, 36, 37
上限, 315
上有界, 43
射, 10
射影, 8, 11
施瓦茨, 323
实根, 255
实数模型, 31, 48
势, 22
等势, 21, 63–65
集 X 的势, 22
连续统的势, 63, 64
势(基数), 21
收敛数列, 67
数
$2\pi R$, 340
π, 43, 132, 248, 336, 341
e, 77, 89, 90, 118–120, 248, 267, 269
超越数, 42, 55, 63, 64
代数数, 42, 55, 63
复数, 238
负数, 36
基本初等函数的导数表, 182
集 X 的基数, 22
实数, 29
素数, 41
无理数, 40, 55, 64
有理数, 40, 41, 55, 62, 63
整数, 40
正数, 36
自然数, 24, 38
自然数集, 27
斐波纳契 (Fibonacci) 数, 92
数 x 的阶, 51
数量积, 385
数列, 68, 72
数列的极限, 67, 68
数轴, 44
双曲、反双曲函数, 180
速度
摆的振动方程, 350
第二宇宙速度, 350
光在这两种介质中的速度, 215
瞬时加速度, 165
瞬时速度, 157, 165
真空中的光速, 48
算子, 10
拉普拉斯算子, 421, 426
平移算子, 12

T

泰勒公式
具积分形式余项的泰勒公式, 422, 454
泰勒公式余项, 197
积分形式泰勒余项, 328
具积分形式余项的泰勒公式, 409
拉格朗日形式的泰勒公式余项, 328
佩亚诺形式, 203
佩亚诺形式余项的泰勒公式, 410
泰勒公式余项, 201
余项的柯西公式, 197
余项的柯西公式和拉格朗日公式, 328
套
闭区间套, 59, 60, 72

开区间套, 61
梯度, 395–397
替换
不定积分中的变量替换, 281
参数的容许替换, 342, 343
积分中的变量替换, 328
欧拉替换, 293
万能替换, 288
条件
多变量函数极值的必要条件, 411
多变量函数可微性的充分条件, 405
极值是否存在和极值不存在的充分条件, 412
可积性的必要条件, 302
可积性的充分条件, 302
条件极值, 467
条件极值的必要条件, 467
同构, 32, 130
同构实现, 32
同阶
f 与 g 是同阶的, 124
同胚
$C^{(p)}$ 类微分同胚, 442
p 级光滑微分同胚, 442, 447, 449
局部微分同胚, 444
微分同胚, 442, 445, 451, 452, 455, 459, 461, 478
最简微分同胚, 451–453
投影, 377
凸性, 218, 220
图像, 19, 149
函数的图像, 230, 234, 417–420
退化临界点, 479
拓扑, 95

W

微分, 158, 179, 387, 388
函数 $f: E \to \mathbb{R}^n$ 在点 $x \in E$ 的微分, 386
微分是函数增量的线性 (主要) 部分, 160
自变量的微分就是它的增量, 160
微分的几何意义, 161
微分法, 171
复合映射的微分法, 393
幂级数的微分法, 251
微分的基本法则, 171
一般的微分法则在复的情形下也都成立, 251
隐函数的微分法, 183
微分算子, 421
尾数, 57
魏尔斯特拉斯最大值定理, 142
位势
力的势, 350
牛顿位势, 349
位置记数法, 50
问题
二体开普勒问题, 154
惠更斯问题, 415
浦丰问题, 351
无穷大
无穷大函数, 122
无穷集, 61
无穷几何级数(或等比级数), 82
无穷小, 122
无穷小函数, 122
无序对, 8, 23
误差, 49, 177, 401
绝对误差, 48, 69, 177
相对误差, 48, 177

X

系数
弹性系数, 270, 275
有效作用系数, 273
下界, 36, 37, 43, 72
下确界, 37
下有界, 37
线性泛函, 314
相对误差, 401

像, 18
向量
　　法向量, 418–420
　　切向量, 419, 420, 465, 476
向量 v 的模, 157
向量场, 399
向量空间
　　空间 $\mathcal{R}[a,b]$, 314
行星轨道, 276
序对, 8, 23
序列, 58, 66, 67
　　有界数列, 74
　　逼近序列, 47
　　不降列, 74
　　不增列, 74
　　常数列, 68
　　单调数列, 74
　　递降列, 74
　　递增列, 74
　　基本列, 72, 241
　　基本数列, 72
　　柯西列, 72, 241
　　上有界列, 74
　　收敛数列, 72
　　下有界数列, 74
　　有界数列, 68
旋轮摆, 350
旋轮线, 350, 364

Y

雅可比行列式, 453
严格单调性, 220
严格局部极大(小)值, 191, 411
严格凸性, 220
一致连续性, 143, 378
因数, 40
　　零因子, 57
　　最大公约数, 55
引理
　　阿达马引理, 423, 454
　　闭区间套引理, 58
　　波尔察诺–魏尔斯特拉斯引理, 77, 81
　　费马引理, 193
　　极限点引理, 60
　　莫尔斯引理, 453, 454
　　有限覆盖引理, 59, 60
隐函数, 428, 441
隐函数定理, 434, 441, 442
映射, 10, 11, 20, 21, 163, 371, 372, 377, 379, 382, 383, 391–393, 397, 442, 443, 459, 462, 463
　　单边逆映射, 20
　　单射, 14, 15, 21
　　复合映射, 14, 380, 393, 394
　　恒等映射, 15, 20
　　互逆映射, 14, 16
　　连续映射, 376, 380
　　满射, 14, 15, 20, 21
　　内射, 20, 21
　　逆映射, 14, 16, 397, 399
　　切映射, 397, 398, 462, 463
　　切映射或导映射, 386
　　上映射, 14
　　双射, 14, 16, 20, 21
　　线性映射, 164, 382–384, 391–393, 398, 399
　　线性映射 $A:\mathbb{R}^m \to \mathbb{R}^n$ 与 $B:\mathbb{R}^n \to \mathbb{R}^k$ 的复合, 383
　　一一映射(又叫双方单值映射), 14
　　一致连续映射, 378
　　映射的复合, 14, 16
　　有界映射, 371
　　右逆映射, 20
　　最终有界映射, 371
　　左逆映射, 20
有界
　　有界集, 43, 60
有界函数, 151
有理函数积分法, 285

有理曲线, 294
有限差, 211
有限集, 62
有限子族覆盖, 59
有效作用系数, 273
域, 31
　　阿基米德序域, 56
　　代数域, 31
　　有序域, 56
元
　　单位元, 30, 33, 36
　　负元, 30, 32, 33
　　极大元, 36, 37
　　极小元, 36, 37
　　零元, 30, 32
　　逆元, 30, 33
　　中性元, 30
　　最大元, 36
　　最小元, 36
原函数, 276, 277, 279–281, 287
　　广义原函数, 325
　　有理函数的原函数, 284, 287
原理
　　阿基米德原理, 43, 44, 56, 61
　　波尔察诺–魏尔斯特拉斯原理, 60, 61
　　博雷尔–勒贝格原理, 59, 61
　　费马原理, 215, 480
　　柯西–康托尔原理, 58, 61
　　上确界原理, 37, 55
　　数学归纳原理, 38
原像, 19
运算
　　乘法对加法有分配性, 31
　　乘法运算, 30
　　初等逻辑运算, 24
　　加法运算, 29
　　结合的, 14
　　微分运算, 171
　　运算 "+" 是交换的, 30
　　运算 "+" 是结合的, 30
　　运算 "•" 是交换的, 30
　　运算 "•" 是结合的, 30

Z

在 E 的极限点 $a \in E$ 处可微, 158
在点 $x \in E$ 处的微分, 159
真假表, 3
真子集, 6
振动, 403
　　简谐振动, 271
　　周期运动, 275
　　周期振动, 271
　　阻尼振动, 273
振子
　　平面振子, 275
　　线性振子, 275
正交规范基底, 385
直积, 8
秩
　　光滑映射 $f: U \to \mathbb{R}^n$ 在点 $x \in U$ 的秩, 446
　　函数组的秩, 457
　　切映射的秩, 446
重排, 243
周期, 271
　　摆的振动周期, 348, 350, 359
　　回转周期, 275
　　粒子的振动周期, 364
　　数学摆的振动周期, 359
　　周期函数, 332
　　周期函数, 函数的周期, 248
逐次逼近法, 15
主值积分, 362
主值意义下的积分, 362
子集, 6, 24
最简分式, 285
坐标, 44, 385
　　笛卡儿坐标, 385, 443
　　第 i 个坐标, 366

点 x 的坐标, 44
极坐标, 443, 444
球坐标, 444
曲线坐标, 418, 444
坐标轴, 44



中文版修订者的话

B. A. 卓里奇的《数学分析》俄文第 1 版 (1981) 是由邝荣雨, 蒋铎, 王昆扬, 钱佩玲和我译成中文的. 俄文原版分 I, II 卷, 中文译本将每一卷分成了两个分册, 高等教育出版社分别于 1987, 1989 和 1994 年出版了前三个分册. 第二卷第二分册因故未能出版. 俄文原著在 1998, 2001 和 2002 年分别出版了第 2、3、4 版. 据作者讲, 德国施普林格出版社已经出版英文版.

卓里奇这套分析教科书的特点在它的序言中已有详细介绍. 这里, 我想引用两位著名数学家, 前苏联科学院院士 A. H. 柯尔莫戈洛夫 (已于 1987 年谢世) 和俄罗斯科学院院士 B. И. 阿诺尔德, 对该书评语中的几段话, 或许能帮助我们更好地理解这套分析教科书的优点.

柯尔莫戈洛夫生前在对这套书的第 1 版的评论中说: 它把叙述的高度严谨性与可读性、充实的内容以及培养研究实际问题的习惯结合起来了.

阿诺尔德对这套书的称赞更具体, 他说: B. A. 卓里奇的教科书是现有供大学数学系、物理系学生用的分析教科书中最成功的. 它与传统分析教科书的重要区别在于, 它一方面更贴近自然科学 (特别是物理学和力学) 的应用, 另一方面, 它比常规的教科书更多地运用了现代数学 (包括代数学、几何学和拓扑学) 的思想和方法. 教程富于思想性, 它清楚地展示了在具体问题研究中现代数学的思想和方法的强大威力. 特别不寻常的是第二卷, 它包括向量分析, 流形上的微分形式理论, 广义函数论和位势理论的引论, 傅里叶级数和傅里叶变换以及渐近展开初步.

阿诺尔德如此称赞卓里奇的书, 与他反对日益加剧的教材专业化趋势有关. 他说: 当今, 像卓里奇这样编写教科书, 应看作是一个创新. 这在古尔沙时代曾经是平常的, 但是, 惹人注意的近半个世纪的教材专业化趋势阉割了分析教程, 留给它的几

乎只是一个个的论证. 现在看来, 重新使分析教程变成有丰富内容的, 显然, 是非常必要的, 这也与大多数大学生未来将从事应用性的工作有关.

两位数学大家的话引起了我们的思想共鸣. 我们从教学和科研工作中也意识到, 提高数学基础课教学质量就是使学生学到符合现代水准的、足够的数学理论, 并培养学生具有应用数学理论分析解决实际问题的习惯和能力. 卓里奇的两卷分析教科书以及他在莫斯科大学力学-数学系的教学实践的着力点也正在于此. 因此, 我们以为, 认真研究一下卓里奇怎样解决上述两方面问题, 必将得到许多有益的启示. 高等教育出版社重版卓里奇这套分析教程的确是件大好事.

去年二、三月高等教育出版社的张小萍老师找到我们, 商量根据卓里奇《数学分析》原著第 4 版修订中译本的问题. 当时, 在第 1 版译者中, 蒋铎先生已经谢世, 除我之外的其他三位同志都有不能脱身的工作. 张小萍老师希望我们来做这项工作, 我们也认为这确实也是一件有意义的事, 就这样, 我硬着头皮接受了这项任务.

卓里奇的《数学分析》第 4 版与其第 1 版比较, 虽然内容、结构、思想、风格依旧, 但具体改动却不少. 每卷末都增加了口试试题和考试大纲; 傅里叶级数和傅里叶变换这一章全部改写了; 对一些关键定理给出了新证明或改写了原来的证明; 充实了与自然科学有关的例子; 重新编排了部分习题, 并增加了一些新习题; 作了更加详细的索引; 许多段落中都有词句改动.

鉴于这种情况, 加之我们的第 1 版译本本身也有一些错误和译得不好的地方, 这次修订采用了以通读俄文原著第 4 版为基础, 逐字逐句对照, 进行修改、重译和补译的方法. 原版中的印刷错误和明显的笔误, 在译文中改正后未作标注. 有个别地方, 修订者认为有必要补充说明, 就写了译者注.

虽然我作了不少努力以减少错误, 但限于语言文字水平和自然科学知识水平, 难免还有不少改错译错之处, 希望有关专家、老师及所有读者不吝赐教, 批评指正.

我感谢邝荣雨、王昆扬、钱佩玲几位同事的信任和支持, 感谢高等教育出版社的张小萍老师在修订过程中对我的帮助以及她为该书顺利出版所做的各种努力. 我还要对书的作者 B. A. 卓里奇教授深深致谢, 1989 年他为第 1 版中译本写了补序. 为这次修订, 他向我们介绍了第 4 版与第 1 版比较所作的主要改动. 还应感谢对译稿加工的徐伯勋、文小西、杨芝馨和郭思旭几位先生, 他们认真细致的工作使修订稿避免了不少错误.

周美珂

2006 年 2 月于北京师范大学